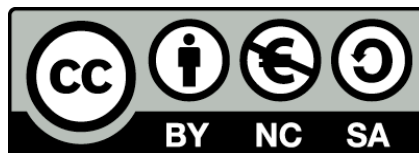UNIVERSITAT DE
BARCELONA

# Pleiotropic effects of candidate genes on autism spectrum disorder and comorbidities: genetics, funcional studies and animal models

Ester Antón Galindo
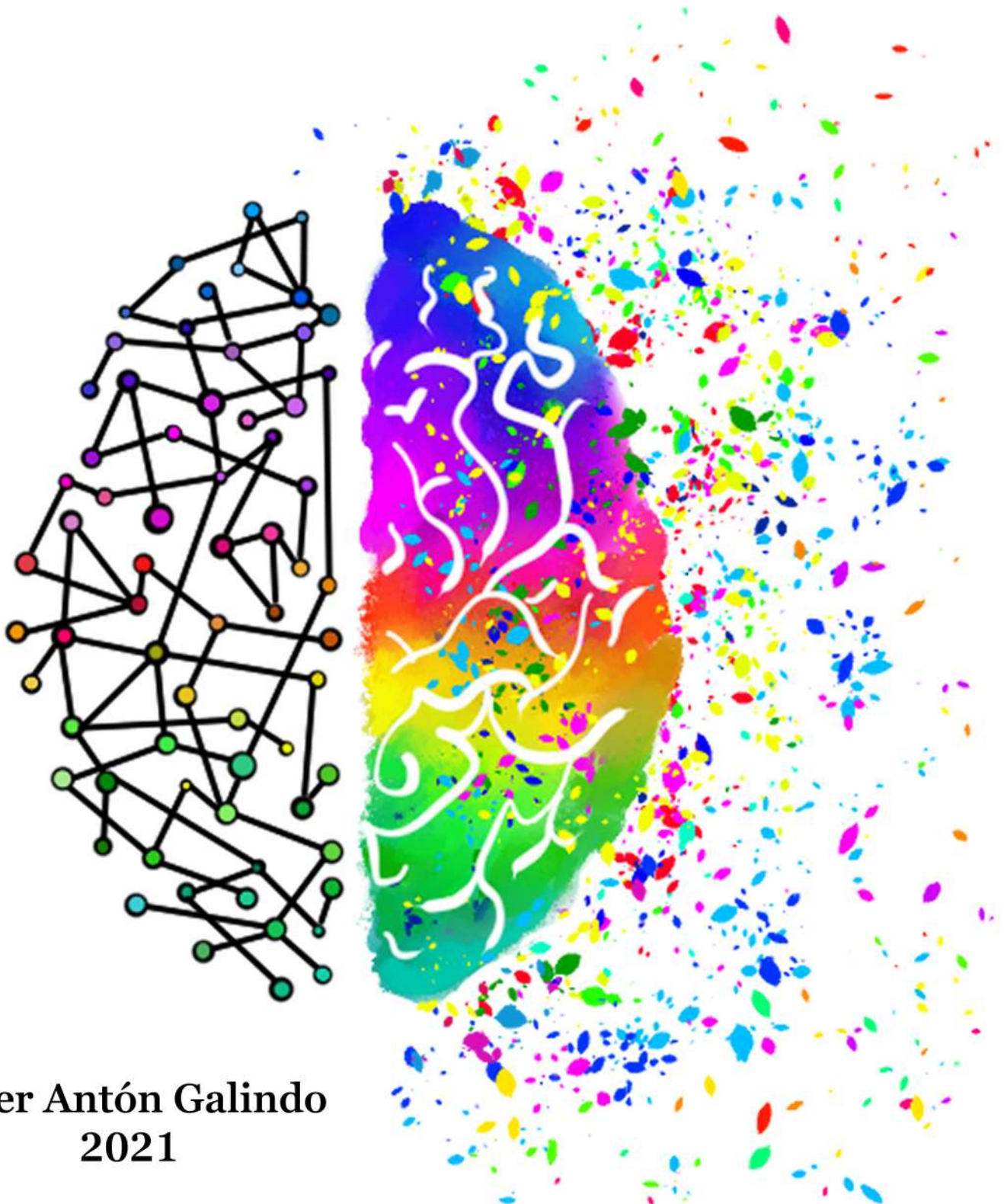
# PLEIOTROPIC EFFECTS OF CANDIDATE GENES

# ON AUTISM SPECTRUM DISORDER AND COMORBIDITIES:

# GENETICS, FUNCTIONAL STUDIES AND ANIMAL MODELS

Ester Antón Galindo
2021

# Pleiotropic effects of candidate genes on autism spectrum disorder and comorbidities: genetics, functional studies and animal models

Doctoral Thesis presented by

**Ester Antón Galindo**

To apply for the title of

**Doctor of the University of Barcelona**

Programme of Genetics

Department of Genetics, Microbiology and Statistics

Doctoral Thesis supervised by **Dr. Noèlia Fernàndez Castillo** and **Dr. Bru Cormand Rifà** at the Department of Genetics, Microbiology and Statistics of the University of Barcelona

Dr. Noèlia Fernàndez Castillo

(codirector)

Dr. Bru Cormand Rifà

(codirector and tutor)

Ester Antón Galindo

Barcelona, June 2021

UNIVERSITAT DE BARCELONA

# AGRADECIMIENTOS

Creo que es muy difícil plasmar en un folio todo lo que una siente y quiere hacer llegar a las personas que quiere. Siento que la tesis ha sido un camino largo, con sus más y sus menos, per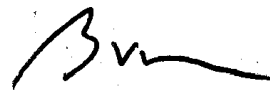o durante el cual he tenido la suerte de contar con personas maravillosas. A pesar de lo complicado que me parece, intentaré transmitir en estas líneas unas pinceladas de todo este agradecimiento.

Allá por el 2016 decidí que quería hacer un doctorado y, por suerte, el destino quiso que Bru y Noe estuviesen buscando en ese momento a alguien. Nos tiramos todos a la piscina, tanto ellos como yo, porque en realidad no nos conocíamos más allá de un par de entrevistas... Parece que salió todo bien a pesar de todo y, cuatro años después, les estoy muy agradecida. Quiero dar las gracias a Bru y Noe, por darme la oportunidad de trabajar con ellos en este proyecto, por todo lo que he aprendido durante estos cuatro años y por todo el apoyo que he sentido por su parte. Noe, durante estos años has sido mucho más que una "jefa" maravillosa, has sido también una amiga y un modelo de científica brillante, gracias por confiar tanto en mí, por esa pasión que pones en todo lo que haces, y por transmitirme esas ganas de aprender y de hacer justicia en la vida.

Quería agradecer también a todas esas personas que he conocido en el laboratorio y que han hecho que ir a trabajar cada día fuese una motivación. Ya desde el momento en que me recibieron todos vestidos de amarillo supe que no me iba a arrepentir de mi decisión: estos años no solo he tenido compañeros de trabajo, sino un grupo amigos. Gracias a mis Neuro-chicas por hacer estos años más amenos, a Bárbara por su alegría, a Judit por su dulzura y paciencia con todo, a Laura por su optimismo incansable, y a Eva por su trabajo tan aplicado y por todas las horas sufridas haciendo nuestros queridísimos BRET. A las más jarras, miembras honoríficas del club del pedal, Nuria y Laura C, por tantas risas fuera y dentro del laboratorio. Laura, gracias por estar siempre ahí, por ser una persona tan importante para mí y alguien con quien sé que puedo contar siempre. Nuria, a pesar del tercer grado que me hiciste nada más llegar te tengo muchísimo aprecio, has sido siempre la cohesión del grupo, gracias por transmitir tanta alegría. To Maja, I am so happy you joined our group! Thank you for these good moments in Barcelona, for all your support when I needed it, and for these tough hours we passed in a cave tracking zebrafish haha A la Meri, porque por suerte nos conocimos antes de tus 25 y tu paso por el lab ha llevado a una preciosa amistad que va a durar mucho, mucho tiempo. A Edgar, porque a pesar de nuestra relación amor-odio sabemos que siempre triunfará el amor (pero no me cortes). A todos los que contribuisteis a esos primeros años increíbles fuera y dentro del laboratorio: Aldo (*aka* el niño de Barbate), Neus, Guillem, Marta, Noe Benetó. A Mónica, por ser el pilar del lab, a los jefes Roser, Kelly, Susanna, Dani, Marina, y a las nuevas y prometedoras incorporaciones, Juan (gracias por todo ese cariño y amor que transmites), Aina, Nerea, Irene, Ainhoa, Estefanía, Mónica C, Mario, que a pesar de llegar en momento complicados mantienen el espíritu de lab.

A todas esas personas con las que he compartido horas de trabajo durante estos años. Thank you, Will, for hosting me in your lab and being always so helpful. Gracias Héctor, por hacerme sentir tan acogida desde mi primer día en Leicester. A Gustavo y demás gente del ICFO, Matteo, Emilio, por todas esas horas de charla grabando pececitos que brillan. A Xavi, porque lo embaucamos en un proyecto sin saber la magnitud de lo que se nos venía encima... y porque sin su ayuda no

habríamos llegado muy lejos. Gracias por tu paciencia y por tu trabajo excelente. Nicola, muchas gracias por tu gran ayuda desinteresada y por hacer que incluso me acabe llevando bien con Python. À Stéphanie, qui m'a si bien acceuilli dans son labo pendant mes deux années de master et qui a contribué à ma passion pour la neuro (malgré les allergies) et à ma décision de continuer dans le monde de la recherche.

Pero estos años no han sido solo un paso por el laboratorio y un proyecto de investigación, han sido mucho más, han sido sobre todo mudarme a Barcelona y vivir momentos preciosos y compartidos.

A mis queridas *gayarricas*, que hacen que nuestro piso sea un lugar bonito en el que vivir. Alicia, gracias por estar siempre ahí, por conocerme tanto, por compartir más de veinte años de caminos paralelos y porque eres una persona muy especial que quiero conservar en mi vida. Indirica, mi cantaora favorita y mi compi de pista de baile, de verdad me alegro mucho de haberte conocido y compartir estos meses juntas. Y por supuesto, no me olvido de nuestro bebé Billy al que no puedo querer más <3

A mis *amebas*, porque lo que unió Biotec hace diez años ya no lo separa nadie. Ana, mi querida francesa a la que echo tantísimo de menos, Marta, Ester e Inés, mis intrépidas montañeras, y Luis, Vicky, gracias por contribuir al "núcleo Barcelona". Aunque los reencuentros son cada vez más difíciles de conseguir ahí seguiremos siempre para lo que sea, unas míticas fiestas de Ejulve, unas excursiones de montañeras o unas tardes de karaoke.

A personas muy importantes que tengo en mi amada Zaragoza, y con las que sé que podré contar siempre. Ander, porque lo que hemos vivido juntos une mucho y que nos quiten lo bailao. Gracias por esas aventuras locas en mil rincones de Europa, y por no perder nunca esa confianza y cariño que nos tenemos. Marina, por tantos años de amistad e historias vividas, y por que se sigan conservando - aunque ya hagamos paseos de señoras mayores. Paula, por nuestra amistad tan duradera que vale millones, gracias por estar siempre ahí. Marta V, por el amor que desprendes y lo que disfruto los cafés de reencuentro. A Pili, Ángel, Espe y Javier, gracias por todo ese cariño y apoyo y por las excursiones compartidas a lugares increíbles de mi querido Pirineo.

A personas que me ha regalado la vida estos últimos años. A mis piñas Sonia y Marina, por la amistad tan bonita que creó ese piso de Carrer de Cartagena. No veo el momento de reencontrarnos en uno de nuestros piñi-trips en cualquier lugar del mundo. A Antonio, por ser mi fiel compi de escalada y fisio de confianza. A América, porque no he conocido mujer más loca y maravillosa al mismo tiempo. A Sergio, por haberte convertido en una persona tan importante en mi vida que espero no perder nunca. Gracias por todo el apoyo que me das, por una crêpe con nutella una tarde de navidad o una peli a distancia en noches de cuarentena. A Bruno, gracias por esas meriendas en Brü, esas noches escuchando rap, y, en definitiva, por ser lo mejor que me pasó en Leicester.

A BPolemic, por hacerme descubrir el increíble mundo del pole y convertirlo en una parte tan importante de mi vida. A Mini, por hacerme sentir tan acogida en la escuela, a Ola, porque lo que sé de pole se lo debo sobre todo a ella, en general a todo el equipo de profes que son unas diosas, y a mis compis de vermús y calistenia, Rut, Iva y Esther. Gracias por toda esa sororidad.

A todos esos amigos que tengo desperdigados por el mundo. A Magui y Carlitos, gracias por acogerme en esa increíble ruta por Argentina que fue un oasis en medio de la tesis. A Brenno, o doutor mais sexy do mundo e meu brasileiro preferido. Alla mia cara Wendy, per questa meravigliosa amicizia che cominciò tanti anni fa. Ti aspetto a Barcellona per festeggiare la nostra gioventù e tutti i momenti che ci rimangono per vivere, non vedo l'ora di rivederti, mi corazón.

Por último, quería dar gracias a esas personas con las que comparto tanta genética, a las que quiero tanto y que han sido siempre un pilar en mi camino. A mi familia, esparcida entre Zaragoza y Cataluña, y en especial a mis tíos Mari y Juan, por hacerme sentir en casa desde mis primeros minutos en Barcelona. A mis yayos Nono y Pilar, por cuidarme, por quererme, por enseñarme tanto. A mi yaya María, por ser una mujer tan fuerte, una guerrera en la vida y enseñarme a afrontarla siempre con una sonrisa. A mi hermano, que lo quiero con todo mi corazón. Gracias por ser un idealista empedernido y no dejar nunca de luchar por lo imposible. A mis padres, por todo el apoyo que me han dado siempre en todo y cada uno de mis proyectos, por enseñarme tanto, sobre todo por enseñarme a ser libre y a querer. Os quiero mucho.

Muchas gracias.

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by impairments in social communication and interaction, as well as repetitive and restricted patterns of behaviour. Although growing evidence supports a main contribution of genetic factors to its neurobiology and hundreds of candidate genes have been identified in recent years, the genetic architecture of the disorder is still not fully understood. Moreover, ASD frequently co-occurs with other developmental and psychiatric disorders, and shared genetic mechanisms are hypothesized to underlie these comorbidities. In this doctoral thesis, we aimed to study the contribution of several candidate genes to ASD and comorbidities. We have focused on the 14-3-3 gene family, *RBFOX1* and the *BEX*/*TCEAL* gene family, performed genetic and functional studies and further characterized the neurobiological effects of their deficiency using animal models.

First, our results suggest a role for the 14-3-3 genes in ASD and schizophrenia (SCZ). Ultra-rare variants in the 14-3-3 genes are enriched in ASD and common and rare variants in the *YWHAE* and *YWHAZ* genes, respectively, are associated with SCZ. We have also reported alterations in the expression of these genes in *postmortem* brains of ASD or SCZ patients. Furthermore, we have demonstrated a loss-of-function effect of a damaging variant in the *YWHAZ* gene present in two siblings with ASD and attention deficit/hyperactivity disorder (ADHD). In addition, we have characterized *ywhaz* expression in zebrafish across development and in adulthood and demonstrated that *ywhaz* depletion causes alterations in behaviour, in neuronal activity and connectivity and in monoamine signalling. The behavioural changes included freezing and were rescued with drug treatments that target monoamine neurotransmission.

Second, we have demonstrated a relevant contribution of common variants in *RBFOX1* to psychiatric disorders and traits. Also, we have shown that a high number of copy number variants (CNVs) spanning *RBFOX1* are reported in patients with psychiatric conditions, the vast majority in patients with ASD or SCZ, and patients with these disorders also show a decreased expression of *RBFOX1* in cortex. Finally, we have used knockout animal models to understand its role in psychiatric disorders, and demonstrated that both mice and zebrafish *RBFOX1*-deficient models present behavioural alterations that can be related to neurodevelopmental disorders such as ASD, ADHD and SCZ.

Third, we found that all *BEX/TCEAL* genes are downregulated in *postmortem* brain regions of ASD and SCZ patients and that rare CNVs spanning several *BEX/TCEAL* genes have been reported in patients with severe neurodevelopmental problems. Furthermore, *Bex3*-deficient mice show anatomical and molecular alterations in brain, an excitatory/inhibitory imbalance and behavioural alterations that can be assimilated to ASD- and SCZ-like symptoms.

El trastorno del espectro autista (TEA) es un trastorno del neurodesarrollo caracterizado por problemas en la comunicación e interacción social, así como patrones restrictivos y repetitivos de comportamiento. El peso de la genética en su etiología es cada vez más evidente, aunque la compleja arquitectura genética del trastorno sigue siendo una incógnita. Además, el diagnóstico de otros trastornos comórbidos es frecuente en pacientes con TEA, por lo que se hipotetiza una base genética común. El objetivo de esta tesis doctoral es elucidar la contribución de varios genes candidatos, concretamente la familia de genes 14-3-3, el gen *RBFOX1* y la familia *BEX/TCEAL*, al TEA y otros trastornos comórbidos, realizando estudios genéticos y funcionales, así como caracterizando los efectos neurobiológicos de su deficiencia en modelos animales.

En primer lugar, nuestros resultados sugieren que variantes ultra-raras en los genes 14-3-3 contribuyen al TEA y que variantes comunes y raras en los genes *YWHAE* y *YWHAZ*, respectivamente, están asociadas a esquizofrenia (SCZ). Además, la expresión de los genes 14-3-3 está alterada en pacientes con TEA o SCZ. Hemos demostrado que una variante patogénica en el gen *YWHAZ* presente en dos hermanos con TEA y trastorno de déficit de atención e hiperactividad (TDAH) provoca una pérdida de función de la proteína. Asimismo, hemos demostrado que la deleción de *ywhaz* produce alteraciones en la actividad y conectividad neuronal, la señalización monoaminérgica y el comportamiento, pudiéndose este último recuperar mediante fármacos.

En segundo lugar, hemos demostrado que variantes comunes en *RBFOX1* están asociadas a diferentes trastornos psiquiátricos y que un número elevado de variantes del número de copias (CNV) afectan a *RBFOX1* en pacientes con trastornos psiquiátricos, siendo especialmente frecuentes en pacientes con TEA o SCZ que, además, presentan una disminución en la expresión de *RBFOX1* en corteza cerebral. Asimismo, hemos usado modelos animales genoanulados para estudiar la implicación de *RBFOX1* en trastornos psiquiátricos, demostrando que tanto el modelo murino como los de pez cebra presentan alteraciones de comportamiento relacionadas con trastornos del neurodesarrollo, como ASD, TDAH y SCZ.

Por último, la expresión de los genes *BEX/TCEAL* está disminuida en regiones cerebrales de pacientes con TEA o SCZ y, además, se han descrito CNVs que abarcan varios genes *BEX/TCEAL* en pacientes con trastornos severos del neurodesarrollo. Los ratones genoanulados para *Bex3*

muestran alteraciones anatómicas y moleculares en cerebro, un desequilibro excitación/inhibición y alteraciones de comportamiento asimilables a síntomas de TEA y SCZ.

# TABLE OF CONTENTS

# ABBREVIATIONS

5-HT: serotonin (5-hydroxytryptamine)

ABIDE: Autism Brain Imaging Date Exchange

ACC: anterior cingulate cortex

ADHD: attention deficit/hyperactivity disorder

ADI-R: Autism Diagnostic Interview-Revised

ADOS-2: Autism Diagnostic Observation Schedule-2nd Edition

ANTs: Advanced Normalization Tools

ASD: autism spectrum disorder

ASL: arterial spin labelling

BAT: behavioural avoidance task

BD: bipolar disorder

*BEX*: Brain Expressed X-Linked (gene family)

BOLD: blood-oxygen level-dependent

BRET: Bioluminescence Resonance Energy Transfer

Ca: cases

CaM: calmodulin

CDD: childhood disintegrative disorder

CD-MA: cross-disorder meta-analysis

CNMF: constrained nonnegative matrix factorization method

CNV: copy number variant

Co: controls

cpGFP: circularly permuted enhanced green fluorescent protein

CRISPR/Cas9: clustered regularly interspaced short palindromic repeats/ CRISPR-associated protein 9

DA: dopamine

dbGaP: database of Genotypes and Phenotypes

DGE: differential gene expression

DLPFC: dorsolateral prefrontal cortex

DNA: deoxyribonucleic acid

DSM: Diagnositic and Statistical Manual of Mental Disorders

DTI: disffusor tensor imaging

*DYRK1A*: Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1A

DZ: dizygotic

EEG: electroencephalography

E/I: excitation/inhibition

EMA: European Medicines Agency

EONDT: early-onset neurological disease trait

ESAT: Early Screening of Autistic Traits

FDA: Food and Drug Administration

FDR: false discovery rate

fMRI: functional magnetic resonance imaging

GABA: gamma-aminobutyric acid

GCaMP: a family of genetically-encoded calcium indicators

GECI: genetically encoded calcium indicator

GEO: Gene Expression Omnibus

GnoMAD: Genome Aggregation Database

GTEX: genotype-tissue expression

GWAS: genome-wide association study

HET: heterozygous

HOM: homozygous

ICD-10: International Statistical Classification of Diseases and Related Health Problems 10th revision

ID: intellectual disability

IGE: idiopathic generalized epilepsies

IQ: intelligence quotient

Indels: insertions and deletions

iPSYCH: integrative Psychiatric Research consortium

KO: knockout

LD: linkage disequilibrium

LFM: light field microscopy

MAF: minor allele frequency

M-CHAT: Modified Checklist for Autism in Toddlers

MDD: major depressive disorder

MEG: magnetoencephalography

MRI: magnetic resonance imaging

MZ: monozygotic

NGS: next generation sequencing

OCD: obsessive-compulsive disorder

OMIM: online mendelian inheritance in man

OR: odd-ratio

PCR: polymerase chain reaction

PDD-NOS: pervasive developmental disorder not otherwise specified

PET: positron emission tomography

PGC: Psychiatric Genomics Consortium

PolyPhen-2: Polymorphism Phenotyping v2

PRS: polygenic risk score

*RBFOX1*: RNA binding protein, fox-1 homolog 1

RDoC: Research Domain Criteria Initiative

RNA: ribonucleic acid

rs-MRI: resting-state MRI

RT: risk taking

SCC: Simons Simplex Collection

SCZ: schizophrenia

SFARI: Simons Foundation Autism Research Initiative

SIFT: Sorting Intolerant From Tolerant

SNVs: single nucleotide variants

*SFN:* stratifin

sMRI: structural magnetic resonance imaging

SNP: single nucleotide polymorphism

SPECT: single-positron emission computed tomography

SSRI: serotonin reuptake inhibitor

*TCEAL*:  transcription elongation factor A (SII)-like (gene family)

TH: tyrosine hydroxylase

TL: Tübingen Long-fin

URVs: ultra-rare variants

UTR: untranslated regions

WES: whole exome sequencing

WGS: whole genome sequencing

WT: wild-type

*YWHAE:* Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Epsilon

*YWHAZ:*  Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Zeta

# INTRODUCTION

# 1. AUTISM SPECTRUM DISORDER

## 1.1. History and definition

The word *autism* comes from the Greek word *autos (*αὐτός*)*, which means "self". The term was coined in 1910, by Paul Eugen Bleuler, a Swiss psychiatrist, to describe morbid self-admiration and withdrawal within self that some schizophrenia patients presented [1]. Then, in 1943, Leo Kanner introduced for the first time the label *early infantile autism* in a systematic report of 11 children presenting social impairment, obsession with objects and a need for sameness [2]. A year later, Hans Asperger, an Austrian paediatrician, published a report about children with a similar pattern of behaviour, characterised by trouble with social interactions and specific obsessive interests [3]. Both clinicians hypothesized a neurobiological origin of autism, and some years later, the first and second editions of the DSM (Diagnostic and Statistical Manual of Mental Disorders) [4,5] defined autism as a behaviour associated with childhood schizophrenia.

However, in the 1950s and 1960s, psychoanalysts such as Bruno Bettelheim, popularized a theory claiming that "refrigerator mothers", as they termed them, caused autism by not loving their children enough [6]. Fortunately, the "refrigerator mother" concept was disproved in the following years by an increase in biological research that found differences in brain function and cognition in "autistic children". It was not until the third edition of the DSM in 1980 [7] that autism was considered a disorder with a separate diagnosis from schizophrenia. This edition defined autism as a disorder with three essential features: lack of interest in people, communication difficulties and inappropriate behaviour to environmental stimuli.

The DSM-IV, released in 1994 and revised in 2000 [8,9], subcategorized autism into five conditions with distinct features: autistic disorder, pervasive developmental disorder/not otherwise specified (PDD-NOS), Asperger's disorder, childhood disintegrative disorder (CDD), and Rett syndrome. This categorization was based on the hypothesis that each of these conditions had a different genetic cause linked to specific problems and treatments. However, the lack of clear borders between all these conditions made an accurate diagnosis of patients difficult. Moreover, genetic studies in autism were not able to point at specific candidate genes for the conditions listed in DSM-IV. For these reasons, the last DSM edition, DSM-5 [10], introduced the term "autism spectrum disorder" (ASD) and eliminated the clinical subtypes defined in the previous version.

## 1.2.  Epidemiology

ASD is one of the most prevalent neurodevelopmental disorders. The 2010 Global Burden of Disease study establishes a prevalence of 1 in 132 individuals [11]. Prevalence estimates have increased significantly in recent years and vary between studies, which is partially attributable to changes in the diagnostic criteria over the years [12] and to the use of different methods of ascertainment. According to a recent review [13], differences in estimates between populations and ethnicities may be affected by socioeconomic disparities, social and structural mechanisms that cause unequal access to diagnosis, treatment and services related to ASD.

Autism is more common in males than in females, with reported ratios ranging from 2:1 to 5:1 and an estimate of 4:1 in the 2010 Global Burden of Disease Study [14,15]. This difference in prevalence may be due to different (epi)genetic and environmental factors, but the mechanisms underlying male vulnerability or female protection are still not fully understood. A recently proposed theory, the multi-hit hypothesis, states that interactions between sex, genes, and environment lead to the male bias in ASD [16]. Moreover, sociocultural factors in the application of the diagnostic criteria may contribute to an inaccurate diagnosis of autism in females [17–20]. Indeed, a study performed by Russell showed a strong gender bias when diagnosing ASD: girls are less likely to be diagnosed even with equally severe symptoms [17].

## 1.3.  Clinical characteristics and diagnostic criteria

Diagnosis of autism is made by trained clinicians based on behavioural traits and the identification of core diagnostic features present in the early developmental period. The current diagnostic criteria for ASD described in DSM-5 [10] are divided into two main clinical areas, summarized in Table 1. To meet diagnostic criteria, a child must have persistent deficits in each of the three areas of social communication and interaction plus at least two of four types of restricted, repetitive behaviours. Once diagnosed, severity specifiers are used to better describe symptomatology.

A number of structured diagnostic interviews and observational assessments for autism exist in order to standardize and facilitate the diagnosis process [21,22]. Nowadays, the most used and validated instruments to diagnose autism are the Autism Diagnostic Interview-Revised (ADI-R) [23] and the Autism Diagnostic Observation Schedule-2nd Edition (ADOS-2) [24]. ADI-R is an interview with questions divided into five different areas: opening questions, communication, social development and play, repetitive and restrictive behaviour, and general behaviour

problems. ADOS-2 consists on a series of structured and semi-structured tasks that evaluate communication, social interaction, play and restricted and repetitive behaviours.

**Table 1. Diagnostic criteria for autism spectrum disorder according to DSM-5** [10]**.**

| Persistent deficits in social communication and social interaction across multiple contexts |
|---|
| Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions. |
| Deficits in nonverbal communicative behaviours used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication. |
| Deficits in developing, maintaining, and understand relationships, ranging, for example, from difficulties adjusting behaviour to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers. |
| Restricted, repetitive patterns of behaviour, interests, or activities |
| Stereotyped or repetitive motor movements, use of objects, or speech (e.g., simple motor stereotypes, lining up toys or flipping objects, echolalia, idiosyncratic phrases). |
| Insistence on sameness, inflexible adherence to routines, or ritualized patterns of verbal or nonverbal behaviour (e.g., extreme distress at small changes, difficulties with transitions, rigid thinking patterns, greeting rituals, need to take same route or eat same food every day). |
| Highly restricted, fixated interests that are abnormal in intensity or focus (e.g., strong attachment to or preoccupation with unusual objects, excessively circumscribed or perseverative interests). |
| Hyper- or hyporeactivity to sensory input or unusual interest in sensory aspects of the environment (e.g. apparent indifference to pain/temperature, adverse response to specific sounds or textures, excessive smelling or touching of objects, visual fascination with lights or movement). |

Finally, the importance of early testing to identify children with autism at a very young age has motivated the development of early diagnosis parent-report instruments such as the Modified Checklist for Autism in Toddlers (M-CHAT) and the Early Screening of Autistic Traits (ESAT) [25–27]. However, the unspecificity and high sensitivity of these instruments may lead to the over-diagnosis of autism [28].

## 1.4. Comorbidities

ASD presents a high level of comorbidity with other pathologies, both in adults and children (Box 1). More than 70% of patients present a comorbid condition: developmental or psychiatric disorders, personality disorders, or general medical pathologies [21] (Figure 1). Common comorbidities are intellectual disability (ID) (45%), epilepsy (8-30%), attention deficit/hyperactivity disorder (ADHD) (28-44%), and aggressive behaviours (up to 68%). Anxiety disorders, depression, sleep disorders, language disorders and gastrointestinal problems are also frequently present in ASD patients [21,29].
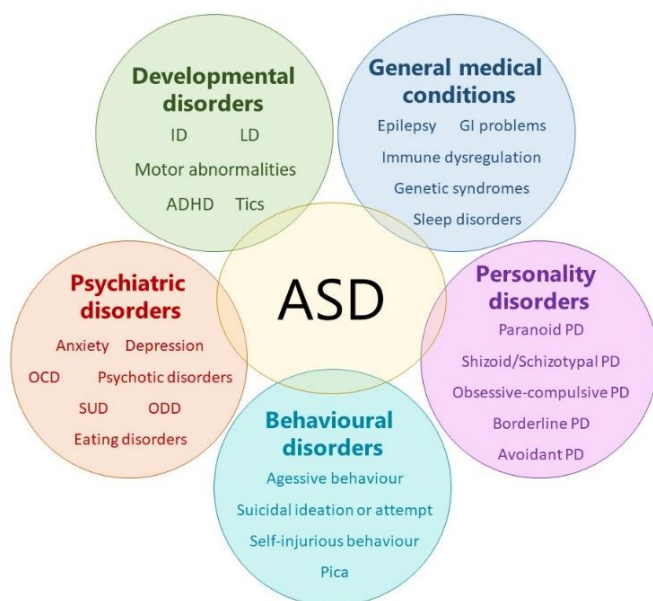


**Figure 1. Comorbid conditions with ASD**, according to Lai, 2014 classification [21]. ADHD, attention-deficit/hyperactivity disorder; GI, gastrointestinal; ID, intellectual disability; LD, language delay; OCD, obsessive-compulsive disorder; ODD, oppositional defiant disorder; PD, personality disorder; SUD, substance use disorder.

---

**BOX 1. Comorbid conditions frequently reported in ASD patients.**

**Intellectual disability (ID)** is characterized by significantly impaired intellectual and adaptive functioning and is defined by an intelligence quotient (IQ) lower than 70. The estimated prevalence of ID in ASD patients is affected by the methodological diagnosis and the definition of intelligence used but is considered to be higher than 30%. A higher IQ predicts a better outcome of ASD patients, as it is directly related to a higher independence during adulthood [21].

**Epilepsy** in autistic patients was first described by Kanner in 1943. Since then, numerous studies have established an increased rate of epilepsy in individuals with autism, especially in patients with comorbid ID [30,31]. Also, the prevalence of electro-encephalogram abnormalities in ASD patients is significantly high (85.8%) [32]. Although the close association between epilepsy and autism is well established, the complex mechanisms that underlie this relationship remain to be fully elucidated [30,31].

---

**Attention-deficit/hyperactivity disorder (ADHD)** is the most common psychiatric comorbidity in individuals with ASD and it has been exhaustively studied in the last years. ADHD is characterized by inattention, impulsivity, and/or hyperactivity that remain relatively persistent over time and result in impairment across multiple domains of life activities. Earlier versions of DSM excluded the possibility of comorbid ASD and ADHD, but since the advent of DSM-5, the ADHD co-diagnosis with ASD is considered. Partially shared symptoms between ADHD and ASD suggest that some neural and genetic mechanisms may be common to these disorders [33]. Indeed, both ADHD and ASD are associated with executive functioning impairments [34]. Also, ASD children presenting with ADHD have more severe impairments in adaptive behaviour compared to children with ASD alone [35].

**Schizophrenia** is a psychiatric disorder with an early onset that has been reported to co-occur in up to 34% of ASD patients. Both disorders include deficits in social interaction and communication and an impaired response to sensory stimuli. Increasing evidence supports an overlap of genetic risk factors and similar underlying mechanisms including alterations in dopamine signalling and brain circuitry [36].

**Aggressive behaviour** is also common in patients with ASD and can take different forms, from minor physical aggression in very young children to verbal aggression in adults. Estimates vary significantly across studies (6-68%), probably due to differences in the definition of aggression and the methodology used. The co-occurrence of ASD and aggressive behaviour is associated with higher functional impairment, more intensive medical interventions, it leads to problems with caregiving and it is a risk factor for later poor outcomes [37].

## 1.5. Physiopathology and neurobiology

Anatomical and functional differences, especially in brain anatomy, neuronal activity and immunity, between autistic patients and controls have been observed in *postmortem*, neuroimaging and electrophysiological studies. Current findings draw autism as a condition resulting from overall brain reorganisation beginning early in development (Figure 2).
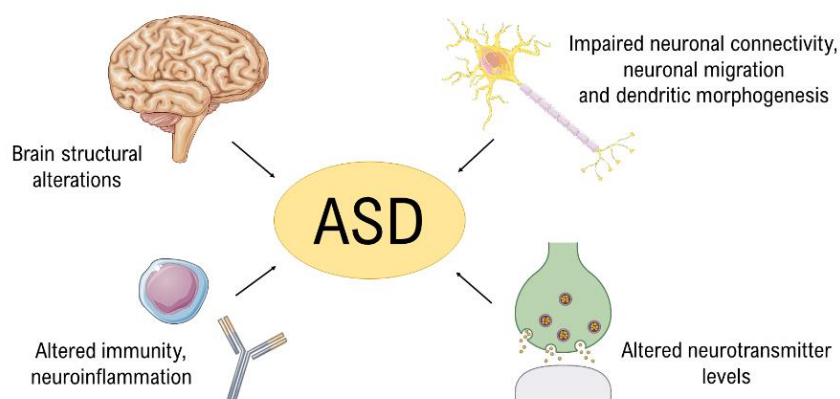


**Figure 2. Altered mechanisms contributing to the physiopathology of ASD.**

**Alterations in brain structures**

Magnetic resonance studies have shown differences in brain anatomy in children with ASD. First, autism is associated with an increased total brain volume during early childhood [38,39]. Indeed, head circumference enlargement, as a measure of brain overgrowth during the first years of life, is common in young children with ASD, especially in low functioning autistics, and macrocephaly is present in around 15% of autistic children [40] (Figure 3). Second, the increase in volume is not uniform across brain areas: imaging studies in autistic patients have reported an enlargement of dorsolateral prefrontal and medial frontal cortex [38,41], basal ganglia [42,43], amygdala [44,45] and hippocampus [46,47], and, conversely, reductions in the size of the corpus callosum [45,48]. However, some of these structural findings have not been consistently reproduced, which may reflect the clinical heterogeneity among ASD patients, but also sample size issues. In addition, several studies have reported a significantly increased number of neurons in the prefrontal cortex of autistic patients and grey matter abnormalities in cortical areas of children and adolescents with ASD [49] (Figure 3).
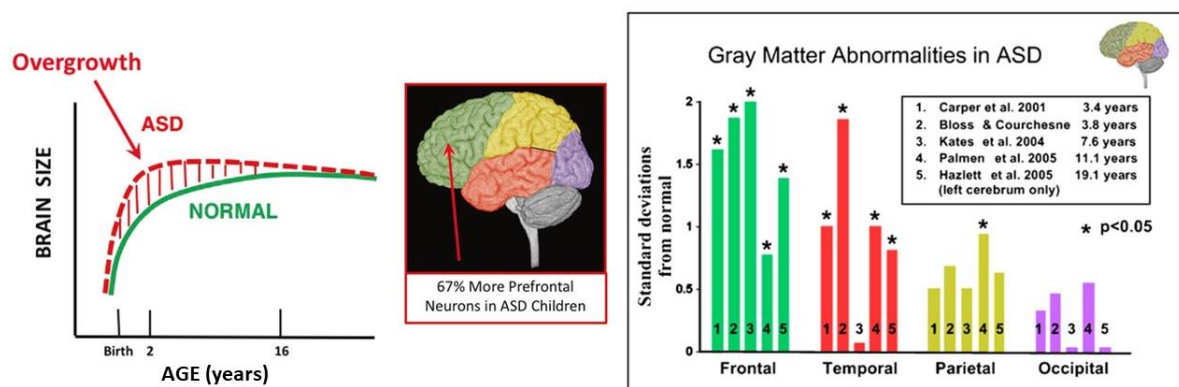


**Figure 3. Early brain overgrowth and grey matter abnormalities reported in the first years of life in children and adolescents with ASD**. Adapted from Courchesne et al., 2007 and 2019 [39,50].

Finally, abnormalities in the cerebellum, a region that plays an important role in cognition and motor function, have been consistently reported in individuals with ASD. MRI studies reported a cerebellar enlargement in autistic children [51] and a reduction in adults [52], a decreased volume of the cerebellar vermis [53–55] - involved in modulating emotion, arousal and sensory response, and a reduced cerebellar grey matter [56,57]. Also, *postmortem* studies found a decreased number and size of Purkinje cells in adolescent and adult patients [58,59].

**Impaired neurogenesis, neuronal migration and dendritic morphogenesis**

The development of neural circuits starts early in embryogenesis and consists of several subsequent steps (Figure 4). A current hypothesis asserts that impaired neurodevelopmental processes, in which shaping and fine-tuning of neuronal circuits occur, may underlie the early-brain overgrowth and brain structural abnormalities observed in ASD patients [39,60]. A defective neurogenesis and neuronal migration during the prenatal period in autistic patients may cause an initial misplacement of neurons and lead to an abnormal growth of brain areas, also called dysplasia, and an altered brain maturation [61]. Indeed, mutations in genes such as *RELN*, with a crucial function in neuronal migration, have been described in ASD patients [62,63]. Also, even if *postmortem* dendrite and spine studies in ASD patients are scarce, an altered dendritic spine density has been reported in these patients [64].
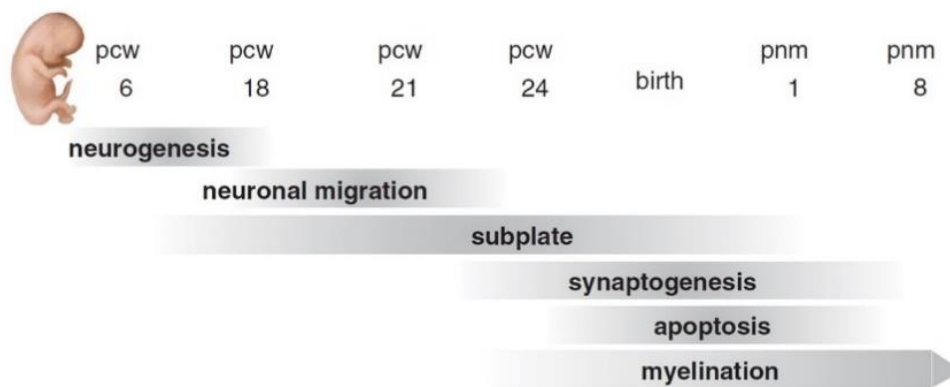


**Figure 4. Steps of neuronal development in humans**. Approximate time points of the major neocortex developmental events in human. pcw, postconceptional weeks; pnm, postnatal months. Adapted from Molnár et al., 2020 [65].

**Impaired neuronal connectivity**

At different developmental periods, spontaneous activity influences or even controls neurogenesis, neuronal migration, synaptogenesis, apoptosis and myelination (Figure 5). Indeed, neuronal communication during early development interacts with genetic programs to establish the mature brain circuitry [65]. In ASD patients, alterations in this spontaneous activity may affect patterning and wiring and lead to an impaired connectivity in relevant brain circuits and a defective lateralization [65,66] (Figure 5). Brain lateralization occurs during typical neurodevelopment [67] and refers to the specialized asymmetry of brain hemispheres in relation to important cognitive and behavioural mechanisms such as memory, language, visuospatial or emotional processing [68–70]. Several studies have reported altered brain asymmetry in ASD patients and a disrupted interhemispheric connectivity and synchronization [71–73].
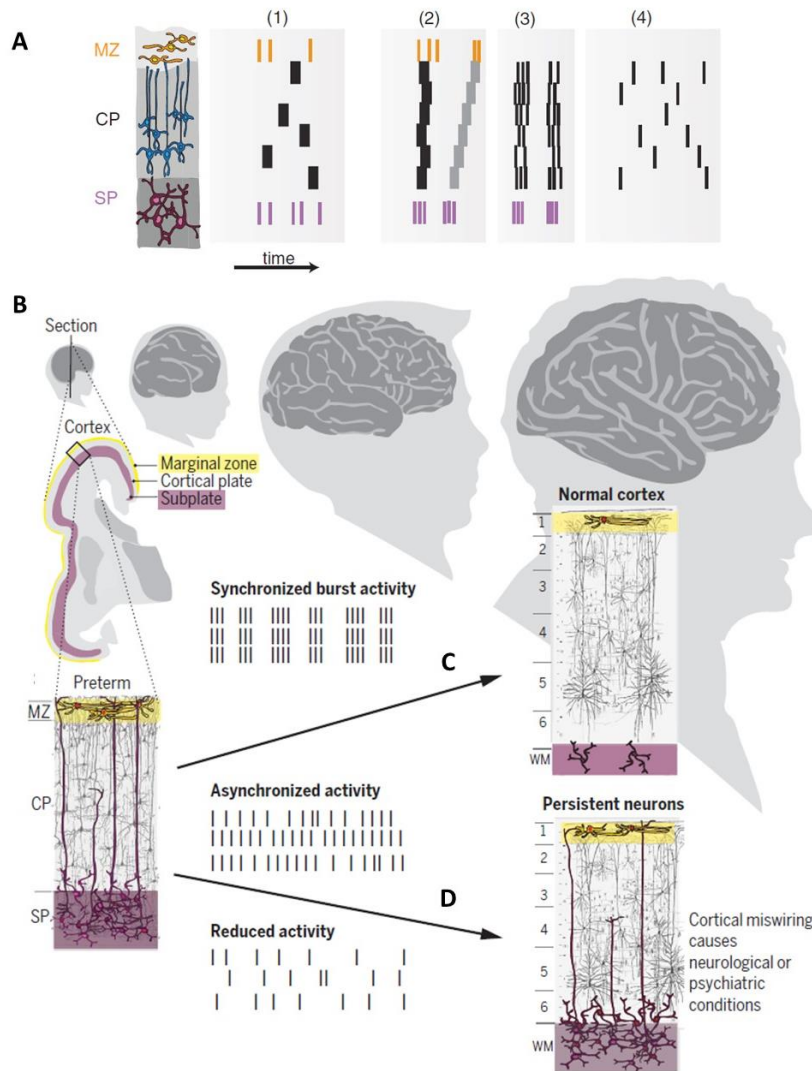
**Figure 5. Spontaneous activity patterns during development modulates cortical wiring. A) Developmental changes in spontaneous cortical activity**: Prenatal cortical circuits are driven by early-generated transient neurons in the marginal zone (MZ) and subplate (SP) before maturation of cortical plate neurons. (1) Cajal-Retzius (yellow) and subplate neurons (purple) discharge fast action potentials at a higher frequency than cortical plate neurons (black); (2) Neurons generate local synchronized activity or propagating activity waves; (3) Discharges become faster and local networks discharge in synchronized bursts, also transient early-generated neurons start to disappear; (4) Adult-like sparse desynchronized activity independent of transient neurons and circuits. **B) Early spontaneous synchronized neuronal activity modulates cortical circuitries.** During development, transformation of early subplate-driven circuits to the adult-like six-layered cortex requires spontaneous synchronized burst activity that controls apoptosis, neurogenesis, neuronal migration and formation and awakening of synapsis. (1) In normal development, most subplate neurons disappear with development, only a few of them survive as interstitial white matter (WM) cells. (2) Genetic or environmental factors can affect neuronal activity during development, and altered activity patterns may disturb subsequent developmental programs leading to miswiring that could cause neurological or psychiatric conditions. MZ, marginal zone; CP, cortical plate; SP, sublpate; WM, white matter. Adapted from Molnár et al., 2020 [65].

In the prefrontal cortex, pyramidal projecting neurons radially aggregate forming minicolumns, in which combinations of GABAergic interneurons modulate pyramidal inputs and outputs. These minicolumns play an integrative sensorimotor role, as they form cortical interlaminar connections that bind sensory-related signals with behaviour-related outputs [74]. Several studies have reported a reduced size of pyramidal neurons, a decreased minicolumnar width due to an increased neuronal density, and a defective inhibitory function in ASD patients [75,76]. This minicolumnar pathology might disable efficient functional connectivity between distant cortical regions and therefore impair cognitive processing in autistic patients [74,75]. Finally, imaging and electrophysiological findings have shown a pattern of overall brain under-connectivity coupled with local over-connectivity within specific regions in ASD patients [77,78]. Importantly, the cerebellum is hyperconnected with sensorimotor cortex but hypoconnected with cognitive areas in ASD patients [79,80].

**Altered immunity, neuroinflammation and gastrointestinal alterations**

There is strong evidence on the association between ASD and dysfunctional immune activity, and on how the effects of immune dysfunction affects neurodevelopment. Indeed, in many cases ASD coexist with immune-based disorders, and a family history of autoimmune disorders is linked with a higher risk of ASD [81,82]. Many studies have found alterations in both innate and adaptive immunity: altered levels of cytokines in the brain tissue, cerebrospinal fluid (CSF) and blood, an increased number of B cells and monocytes, a decreased number of CD4+ T cells, and a higher number of natural killer cells yet these cells show a diminished response to various signals [82,83]. Finally, many ASD patients present gastrointestinal complications together with immune dysregulations, and increasing evidence points to alterations in the microbiota gut-brain axis as a key factor leading to neuroinflammation in autistic patients [84–86]. (Figure 6)
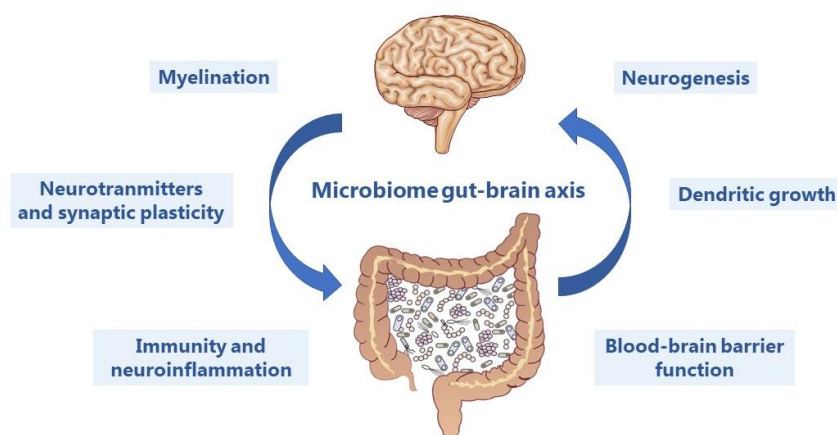


**Figure 6. Gut microbiota impact on brain functioning**. Functions found to be impaired due to gut microbiota alterations. Adapted from Cryan et al., 2019 [86].

INTRODUCTION

**Alterations in neurotransmitter systems**

Growing evidence from molecular, genetic, imaging studies and animal models indicates an important implication of neurotransmitter systems in ASD, especially serotonergic, dopaminergic, GABAergic and glutamatergic systems [87] (Table 2).

**Table 2. Main findings in autism neurochemistry**. Adapted from Marotta et al., 2020 [87].

| Neurotransmitter | Imbalance | Genes | Animal models | Pharmacological approach |
|---|---|---|---|---|
| Serotonin | ↓ 5-HT2A, 5-HT1A binding<br><br>↑brain and blood | SLC6A4 | SERT Ala56 mice<br><br>Slc6a4$^{+/-}$ mice | Selective serotonin reuptake inhibitor |
| Dopamine | ↓ prefrontal cortex<br><br>Dysregulation of mesocorticolimbic and nigrostriatal circuit | SLC6A3, SHANK3, DRD3, DRD4 | Stereotypic deer mice<br><br>DAT T356M$^{+/-}$ | Dopamine receptor blockers |
| GABA | ↓ motor, visual, auditory, somatosensory cortex<br><br>↑blood | MECP2, GABRA5, GABRG3, GABRB3 | Viaat-Cre mice | Arbaclofen, acamprosate, bumetanide, valproate |
| Glutamate | ↓ striatum<br><br>↑blood | SHANK, NGLN3, NLGN4, UBE3A, GRIN2A, GRIN2B, CDKL5 | Nlgn3 KO mice<br><br>Shan3 KO mice<br><br>Shank2 KO mice<br><br>VPA-mice<br><br>Cdkl5 KO mice<br><br>IB2 KO mice | D-cycloserine, memantine, amantadine, mGLuR5-antagonists |

Serotonin (5-HT) modulates essential developmental processes and brain functions such as memory, learning, sleep and mood (Figure 7) and numerous studies have shown its involvement in the aetiology of ASD. Approximately one third of autistic patients show elevated serotonin (5-HT) blood levels [88] and alterations in the serotoninergic system have been reported in ASD patients. These alterations include reduced receptor-binding density of 5-HT$_{1A}$ and 5-HT$_{2A}$ receptors in *postmortem* thalamus, posterior cingulate cortex and fusiform gyrus of ASD patients [89,90], and reduced affinity of the 5-HT transporter SERT in both children and young adults with ASD [91,92]. In addition, polymorphisms in the *SLC6A4* gene, encoding for the 5-HT transporter, have been related to autism [93,94]. Interestingly, a small-scale study reported that the administration of fluoxetine, a selective serotonin reuptake inhibitor (SSRI), leads to a

reduction in repetitive behaviours and an overall improvement of symptoms in adult ASD patients [95]. However, other SSRI such as citalopram are ineffective to treat repetitive behaviours in children with ASD [96].
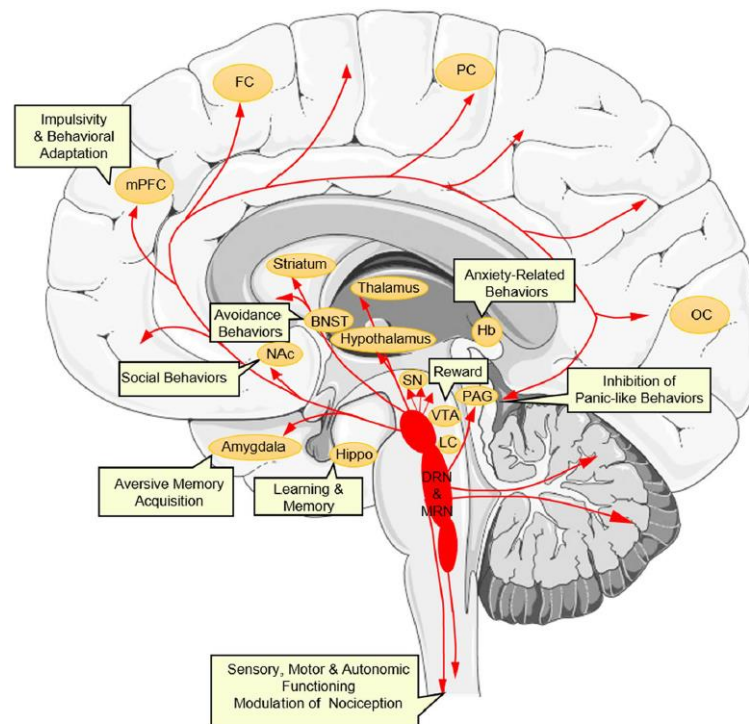


**Figure 7. Serotoninergic pathways in the brain**. Main serotoninergic nuclei and projections are represented in red. Yellow circles represent the brain structures where these projections converge. Main cognitive and behavioural functions in which 5-HT is involved are described. BNST, bed nucleus of the stria terminalis; DRN, dorsal raphe nuclei; FC, frontal cortex; Hb, habenula; Hippo, hippocampus; LC, locus coeruleus; mPFC, medial prefrontal cortex; MRN, median raphe nuclei; NAc, nucleus accumbens; OC, occipital cortex; PAG, periaqueductal gray; PC, parietal cortex; SN, substancia nigra; VTA, ventral tegmental area. Pourhamzeh et al., [97].

Dopamine (DA) plays an important role in motor control, social cognition and reward. Early studies found evidence pointing to the implication of DA dysfunction in autism; in particular, in ASD patients DA hydroxylase activity was found reduced in plasma [98], DA levels were decreased in isolated platelets [99] and a low dopaminergic activity was reported in prefrontal cortex [100]. Now, increasing evidence support the theory that DA imbalance in specific brain regions may contribute to autistic behaviours [101,102] (Figure 8). On the one hand, social deficits in autism may be determined by a dysfunction of the mesocorticolimbic circuit, implicated in social motivation and reward. On the other hand, a dysfunction in nigrostriatal circuits, involved in controlling goal-directed motor behaviour, may contribute to stereotyped behaviours [101,102]. Interestingly, the DA receptors blockers risperidone and aripiprazole are drugs approved by the FDA (Food and Drug Administration) and EMA (European Medicines Agency) to treat agitation and irritability symptoms associated with ASD [22].
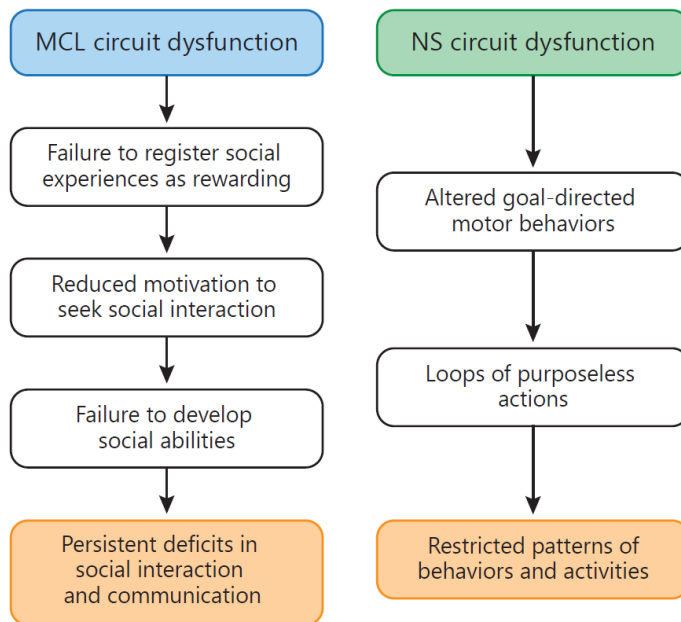
**Figure 8. Dopaminergic pathways suggested to be involved in autistic core features**. MCL, mesocorticolimbic; NS, nigrostriatal. Pavăl 2017 [102].

Finally, balanced neural circuits are important for proper social and emotional behaviour, language processing and cognition. Gamma-aminobutyric acid (GABA) is the main inhibitory neurotransmitter in the central nervous system (CNS) and plays an important role in the development of functional brain pathways. Altered levels of GABA have been observed in plasma and *postmortem* brains of ASD patients, and the excitation/inhibition imbalance theory suggests that alterations in GABAergic signalling could contribute to a loss of balance in neural circuits and lead to an impaired cellular information processing in autistic patients [103].

# 2. GENETICS OF AUTISM SPECTRUM DISORDER

## 2.1. Relevance of genetics factors: heritability studies

Autism is a multi-factorial disorder with genetic, environmental and epigenetic factors contributing to its pathogenesis. However, despite the multi-factorial aetiology of the disorder, epidemiological evidence from family and twin studies supports a strong genetic component underlying ASD and, although there is a high heterogeneity among studies, a recent meta-analysis estimated heritability to range between 64 and 91% [104] (Figure 9). Heritability stands for the degree of variation of a phenotypic trait in a population that is due to inherited genetic variation among individuals in that population. Early estimates of ASD heritability, obtained from family and twin studies, helped to understand the degree to which genetic factors contribute to

ASD susceptibility. Family studies determine the probability that a relative of the studied proband presents the disorder, compared to the prevalence of the disorder in the general population.

The first family studies of ASD estimated a 3% heritability of ASD in siblings, much higher than in the general population, and found that proband relatives were more frequently affected by phenotypic traits related to autism, pointing to the important contribution of genetic factors to ASD aetiology [105,106]. In 1977, the first twin study of autism reported a cohort of 11 monozygotic (MZ) twins and 10 dizygotic (DZ) twins [107]. This study showed a 36% concordance in ASD diagnosis in MZ twins and 0% in DZ pairs when a strict ASD diagnosis was considered, and 92% and 10%, respectively, when a "broader autism phenotype" was used. In the following years, other twin studies were conducted and reported a concordance of 70-90% for ASD in MZ and of around 10% in DZ twins [108–111].
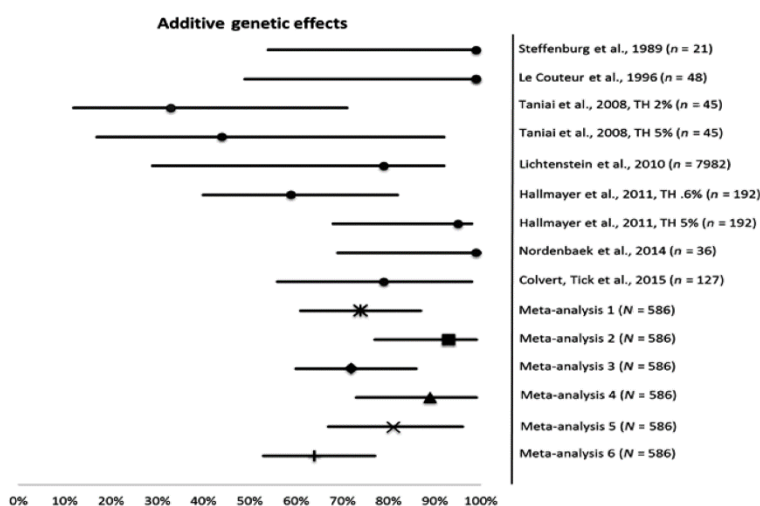


**Figure 9. Heritability estimates of ASD calculated for several studies and meta-analyses.** Forest plots representing additive genetic effects calculated for each study individually and meta-analysis estimates using 6 different configurations. Horizontal lines represent the 95% confidence intervals. Adapted from Tick et al., 2016 [104].

Nevertheless, we should be aware of the fact that estimates extrapolated from these classical twin and family studies might be affected by questionable assumptions in the heritability model [112]. First, the liability model used in these studies does not include known and potential sources of Mendelian and non-Mendelian genetic risk (non-additive effects, epistatis, mitochondrial effects, gene-environment interactions and correlations, etc.). Moreover, these studies assume a small effect of gene-environment interactions: they consider that common environment is shared to the same extent between MZ and DZ twins, and that disparity in the developmental environment of twins and general population has no influence. In recent years, three large population-based studies with a more precise and complete design have determined a refined estimate of 50% heritability [111,113,114], and when all twin studies until date are taken into account, concordance for ASD is 45% for MZ and 16% for DZ twins [115].

In summary, all this evidence from family and twin studies shows that most of ASD risk is driven by genetic factors, yet these studies do not provide information on the specific genetic variation that shapes ASD liability. Indeed, the genetic architecture of ASD is highly heterogeneous and includes different types of variants: chromosomal abnormalities, copy number variants (CNVs), insertions and deletions (indels) and single nucleotide variants (SNVs). These forms of genetic variation differ in frequency (common and rare variants), model of inheritance (autosomal inherited, X-linked and *de novo*), and mode of action (additive, recessive, dominant and hemizygous), all contributing to ASD risk.

## 2.2. Genetic architecture of ASD

**From karyotype and linkage studies to candidate genes**

Cytogenetic studies were the first studies to associate genetic variation with ASD (Figure 10). Early karyotype studies shed light on which genome regions were involved in ASD [116] and, later on, linkage studies examined the co-segregation of chromosomal regions with ASD in families with several affected members. These linkage studies found susceptibility loci in chromosome regions such as 2q, 3q, 7q, 15q and 16p [117–123]. However, the low resolution of these approaches made impossible to associate specific genes with ASD.
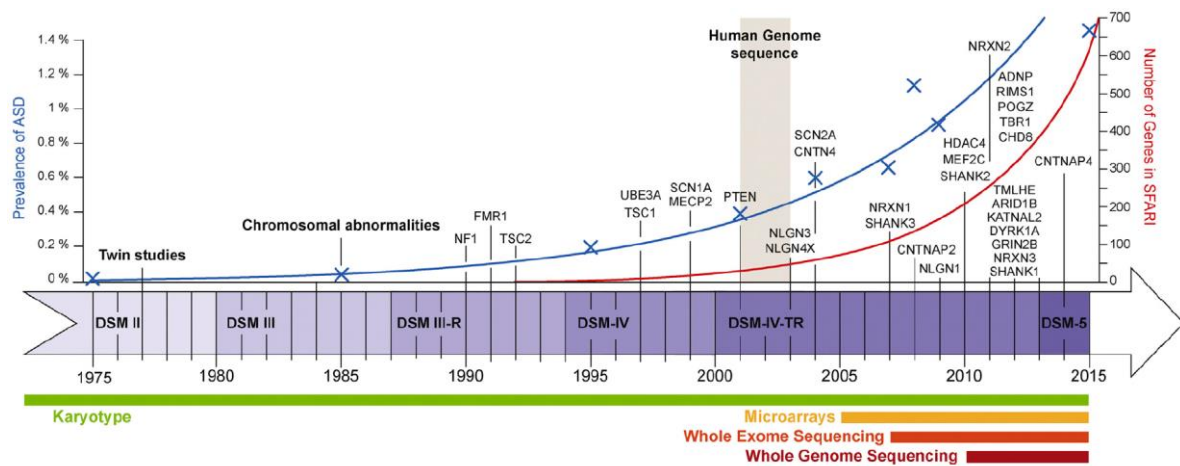


**Figure 10. History of genetic findings in ASD from 1975 to 2015.** The blue line represents the estimates of prevalence of ASD in general population (data from the Center for Disease Control and Prevention) of ASD and the red line the number of genes related to ASD reported in SFARI (gene.sfari.org). Bourgeron, 2016 [115].

Sequencing of the human genome [124,125], together with improvements in methods for detecting variants in PCR-amplified DNA, allowed single genes to be screened for genetic variation in hundreds of individuals. The first approach followed to associate single genes with

ASD relied on the selection of specific candidate genes based on data generated by functional or genetic association studies, according to expression patterns, involvement in neurodevelopment or relevant functions in the CNS. These first studies pointed to genes such *RELN*, *ARX*, *MeCP2*, *NLGN3*, *NLGN4*, *TSC2*, or *UBE3A* [126–131]. However, these "candidate gene" approaches sometimes analysed the wrong genes, lacked enough sample size, and suffered from publication bias.

**Copy number variations**

Copy number variations (CNVs) are submicroscopic structural variants in chromosomes that include duplications, deletions, translocations, and inversions, sometimes covering several kilobases and generally caused by recombination errors during meiosis. The development of genotyping microarray techniques, together with the availability of a reference sequence from the human genome, lead to the discovery of CNVs in 2004 [132,133] (Figure 10). Since then, numerous studies have analysed the contribution of CNVs, inherited or *de novo*, to phenotypic variation among individuals and to disorders such as ASD. These microarray approaches have a much higher resolution than traditional karyotype studies, which were only able to detect 5-10 Mb alterations in chromosomes.

In 2007, the first CNV study performed in ASD patients found a higher load of rare CNVs in autistic individuals compared to controls, underscoring the role of these variants in ASD pathology [134]. An enrichment of rare CNVs in ASD patients was confirmed in later studies [135–142], and it is estimated that around 4-10% of ASD patients present *de novo* CNVs, a higher percentage than in the general population [143]. Recurring CNVs, both inherited and *de novo*, have been associated with ASD in genes such as *NGLN3*, *NLGN4X*, *SHANK3*, *ASTN2*, *PTCHD1*, *NRX1N*, *CNTN4*, *NLGN1* or *SHANK1*, all involved in brain functions, and there are predicted to be 130-234 loci affected by CNVs related to ASD [144,145].

**Next generation sequencing and identification of rare variants**

Next generation sequencing (NGS) revolutionized genetic research and made possible to investigate ASD on a genome-wide level (Figure 10). Whole-genome sequencing (WGS) and whole-exome sequencing (WES) are hypothesis-free approaches, therefore not biased by a previous selection of candidate genes, whose aim is to detect genetic variants associated with the disorder in patients and relatives.

Rare variants - with a minor allele frequency (MAF) of less than 1% - can be inherited from unaffected parents or arise as *de novo* mutations during the meiotic divisions of gametogenesis.

The first WES studies in ASD families focused on identifying *de novo* rare damaging variants [146–149], giving that previous CNVs studies had pointed to the important contribution of *de novo* variants to ASD aetiology. Subsequent research also explored the contribution of inherited variants that are recessive or with incomplete penetrance [150–155]. These first studies generated a list of candidate genes where recurrent *de novo* mutations were found, such as *BRCA2*, *FAT1* and *KCNMA1* [147]; *SCN2A*, *KATNAL2* and *CDH8* [149]; *CHD8*, *DYRK1A*, *GRIN2B*, *PTEN*, *TBL1XR1* and *TBR1* [148]. Interestingly, several of these WES studies showed a correlation between a higher number of *de novo* "gene-truncating variants" (nonsense variants, *indels* leading to frameshift, or canonical splice site variants) and lower IQ in patients with ASD [148,156,157]. A similar association had already been described in previous CNVs studies, which found a higher number of CNVs in ASD patients with ID (IQ<70) [158]. Nevertheless, even if *de novo* highly damaging variants are considerably overrepresented in ASD patients with comorbid ID, they cannot fully explain the disorder in all these patients: the rate of *de novo* truncating variants per ASD proband is quite low (0.2) and it is unlikely that they would be fully penetrant [157]. These variants may explain severe phenotypes, but they are responsible for only a small fraction of ASD patients with ID.

The cost of NGS is continuously decreasing and, consequently, short-read WGS is starting to replace WES as a method to identify genetic factors associated with ASD [159]. WGS allows to identify not only rare coding and non-coding variants, but also small CNVs, genome translocations and inversions, and offers even more coverage across the genome than previous genomic methods, resulting in the discovery of 10-20% more coding variants than WES [160–164].

To date, dozens of large-scale genetic studies has been conducted on ASD patients and their families, leading to hundreds of risk genes being identified. Interestingly, the majority of these genes code for proteins involved in synapse formation or in transcriptional regulation and chromatin-remodelling pathways [165]. In order to gather information about all ASD risk genes and facilitate research, the Simons Foundation Autism Research Initiative (SFARI, www.sfari.org) database was created to provide updated information on genes and CNVs associated with ASD based on published studies (Figure 10).

**Genome-wide association studies and common variants**

Genome-wide association studies (GWASs) emerged in the early 2000's and have been used since then to study complex diseases. An important consideration is that GWASs are hypothesis-free, which means that they are not biased by previous knowledge of location or function of

known genes. The aim of these association studies is to identify genetic variants of risk across the whole genome by comparing allele frequencies between patients and a control population. Individual genomes differ from each other in genetic variants called single nucleotide polymorphisms (SNPs). The majority of these variants are common and occur at least in 1% of the population, but some of them may increase the risk of developing complex polygenic diseases. Indeed, several studies point to common genetic variants acting additively as a major source of risk for ASD [111,114].

While linkage studies were able to detect only highly penetrant variants, GWASs enable detection of SNPs with small additive effects on the phenotype. The first GWAS in ASD was performed in 2006 using microsatellites in a very small sample from Faroe islands and detected associations with 2q, 3p, 6q, 15q, 16p and 18q [166]. Two years later, Arking et al. used 500,000 SNPs in multiplex families with a strict autism diagnosis, but no signal overcame the significance threshold [167]. In the following years, more GWASs were performed, but all of them were underpowered due to the weak effect of individual SNPs and insufficient sample size, and then failed to identify reproducible SNP associations [168–173]. Indeed, these studies found that common variation in the form of SNPs had a very weak effect when a given SNP is considered individually.

The last GWAS to date doubled the discovery sample size from previous studies (more than 18,000 cases) and reported for the first time common risk variants robustly associated with ASD [174]. Five significant loci were defined in ASD alone, and seven additional sites were suggested at a stricter threshold by using GWAS results from three correlated phenotypes (schizophrenia, depression and educational attainment). Several genes located in the identified *loci*, such as *PTBP2*, *CADPS* and *KMT2E*, had been previously related to ASD in genetic studies of *de novo* and rare variants. In addition, this study reported that the polygenic contribution of common variants may be higher in high-functioning ASD patients than in patients with comorbid ID. (Table 3)

**Table 3. Loci reaching genome-wide significance in ASD scans and MTAG analysis with three correlated phenotypes (schizophrenia, major depression and educational attainment) in the last GWAS.** Adapted from Grove et al., 2019 [174].

| Index variant | Chr. | BP | Analysis | p-value | Nearest genes |
|---|---|---|---|---|---|
| rs910805 | 20 | 21248116 | ASD | $2.04 \times 10^{-9}$ | *KIZ, XRN2, NKX2-2, NKX2-4* |
| rs10099100 | 8 | 10576775 | ASD | $1.07 \times 10^{-8}$ | *C8orf74, SOX7, PINX1* |
| rs201910565 | 1 | 96561801 | Comb ASD | $2.48 \times 10^{-8}$ | *LOC102723661, PTBP2* |
| rs71190156 | 20 | 14836243 | ASD | $2.75 \times 10^{-8}$ | *MACROD2* |
| rs111931861 | 7 | 104744219 | Comb ASD | $3.53 \times 10^{-8}$ | *KMT2E, SRPK2* |
| rs2388334 | 6 | 98591622 | ASD-Edu | $3.34 \times 10^{-12}$ | *MMS22L, POU3F2* |
| rs325506 | 5 | 104012303 | ASD-MD | $3.26 \times 10^{-11}$ | *NUD12* |
| rs11787216 | 8 | 142615222 | ASD-Edu | $1.99 \times 10^{-9}$ | *MROH5* |
| rs1452075 | 3 | 62481063 | ASD-Edu | $3.17 \times 10^{-9}$ | *CADPS* |
| rs1620977 | 1 | 72729142 | ASD-MD | $6.66 \times 10^{-9}$ | *NEGR1* |
| rs10149470 | 14 | 104017953 | ASD-MD | $8.52 \times 10^{-9}$ | *MARK3, CKB, TRMT61A, BAG5, APOPT1, KLC1, XRCC3* |
| rs16854048 | 4 | 42123728 | ASD-MD | $1.29 \times 10^{-8}$ | *SLC30A9, BEND4, TMEM33, DCAF4L1* |

In grey, loci reaching genome-wide significance in the analysis of the ASD phenotype alone. In white, additional loci reaching genome-wide significance in the three MTAG analysis performed in the study, with schizophrenia (SCZ), educational attainment (Edu) and major depression (MD). The "analysis" column indicates if the p-value comes from the original scan (ASD), from the combined analysis with the follow-up sample (Comb ASD) or from any of the MTAG analysis (ASD-Edu or ASD-MD). Independent loci are defined to have $r^2<0.1$ and distance >400kb. Chr, chromosome; BP, chromosomal position in base pairs from the telomere of the short arm; "nearest genes", list of the nearest genes from within 50kb of the region spanned by all SNPs with $r^2 \geq 0.6$ to the index variant.

**Epigenetic and environmental factors**

Beyond genetic factors, there is evidence for a role of non-genetic sources of risk in ASD. Epigenetic alterations are likely to contribute to autism, but all DNA methylation studies performed so far, including epigenetic-wide association studies (EWAS) have used small samples. As a result, how epigenetic mechanisms are involved in ASD is still obscure [175,176]. Environmental risk factors for autism have also been reported in recent reviews and studies: advanced parental and maternal age, birth trauma, maternal obesity, a short interval between pregnancies, gestational diabetes mellitus and valproate use during pregnancy. Importantly, these factors cannot be considered causal but could contribute to the risk of autism through several complex underlying mechanisms, such as genetic and epigenetic effects, inflammation and oxidative stress, or hypoxic and ischemic damage [177–179].

---

On one side of the genetic architecture of ASD are the monogenic disorders, in which a single highly damaging gene mutation or a CNV are the major contributors to risk. Notably, in these cases, gene-truncating variants or gene-disrupting CNVs might lead to severe phenotypes such as ID. On the other side of the spectrum is polygenic risk, conferred by the additive effect of common variants, that might shape a genetic background providing susceptibility to the disorder, perhaps in combination with environmental triggers. Inherited common variants with small effect size might therefore increase the risk for ASD development and define a substrate of susceptibility. However, they may not be sufficient for the disorder, and additional variants with moderate or large effect size together with non-genetic factors would then act on this substrate to surpass the liability threshold that results in ASD. The gap between these two extremes spans a broad spectrum in which the contribution of both common and rare *de novo* or inherited variants leads to the disorder [180–182] (Figure 11B). This potential model is supported by the results of WES studies and the latest GWAS, which reported a more important polygenic contribution of common variants in high functioning ASD patients than in the ones with comorbid ID.

**From a polygenic to an omnigenic model**

Analysis of rare variants in ASD cohorts showed that the clinical severity of known pathogenic mutations is influenced by additional rare variants present in the genetic background [141]. Also, the clinical outcome of individuals that carry a rare variant of large effect can also be influenced by the background of common polygenic variation [184,185]. These results point to a polygenic inheritance, in which risk for ASD is determined by the additive contribution of variants in a high number of genes [180], possibly tens or more than one hundred genes per individual.

The advent of GWASs and WES has helped to better understand the genetic basis of complex traits, and has shown that polygenicity is the rule in psychiatric disorders [186], as it seems to be in ASD. More precisely, our current understanding of ASD genetics is compatible with the "omnigenic model" first described by Boyle [187]. In this model, genes are subdivided into core genes, which affect disease risk directly, and peripheral genes, that can only affect risk indirectly through trans-regulatory effects on core genes. This theory postulates that most of the heritability of complex traits is explained by the effects of peripheral genes, propagated through genetic regulatory networks, and that genetic variation in core genes explains only a small part of the overall heritability [188]. "Polygenic" describes the involvement of a high number of variants across the genome in a specific trait, while "omnigenic" would be a special case in the polygenic spectrum in which heritability is mainly driven by peripheral genes that trans-regulate

core genes. However, this distinction of "core" and "peripheral" genes has been challenged in recent reports [189].

**Pleiotropic effects in ASD**

A recurrent term has emerged in recent years when considering genetics of complex traits: pleiotropy [190,191]. Pleiotropy occurs when a genetic variant or a gene has effects on multiple phenotypes, and the omnigenic model of complex traits predicts that network pleiotropy may contribute to genetic correlations among disease phenotypes. These pleiotropic effects would contribute to the high comorbidity of ASD with other psychiatric and medical conditions. Indeed, two recent GWASs cross-disorder meta-analyses highlighted significant genetic correlations between psychiatric disorders, suggesting a complex partially shared genetic structure underlying these disorders [192,193]. In Lee et al. meta-analysis, nearly 75% of the genome-wide significant SNPs were associated with more than one disorder, and ASD was implicated in 36% of the pleiotropic loci, showing the strongest genetic correlations with major depressive disorder (MDD, 45%), ADHD (37%) and schizophrenia (SCZ, 22%). Also, in this analysis the most pleiotropic locus was located in the *DCC* gene and showed association to the eight psychiatric disorders analysed, and the second most pleiotropic locus was identified in the *RBFOX1* gene and showed association to all except one disorder (anorexia) [192]. Finally, the largest GWAS in ASD to date reported several genetic correlations between ASD and other traits, including MDD, SCZ and ADHD [174] (Figure 12). Some examples of single-gene pleiotropic effects would be *SHANK2*, contributing to both ASD and ADHD [194], and neurexins (NRXN1), neuroligins (NLGN4) and *SHANK3* contributing to both ASD and schizophrenia [195].
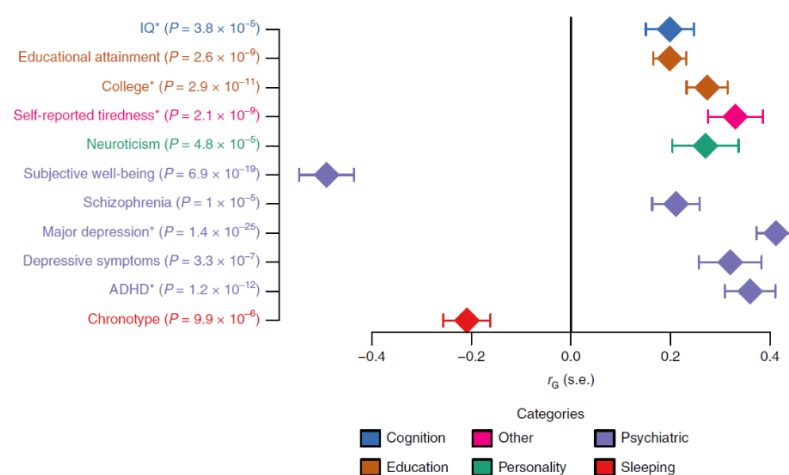


**Figure 12. Genetic correlation between ASD and other traits.** Significant genetic correlations after Bonferroni correction. ADHD, attention deficit/hyperactivity disorder; IQ, intelligence quotient. Adapted from Grove et al., 2019 [174].

**The continuum model of ASD**

The majority of studies in ASD so far have compared cases to controls without considering the possibility of intermediate outcomes. However, an emerging viewpoint suggest ASD as a continuum phenotype with a normal distribution of autistic traits in the general population, where a severe diagnosis (i.e. being 'labelled' as a patient) is at one tail of the distribution [196–199]. Indeed, there is a high variability in social communication and interaction in the general population and subthreshold autistic traits have been observed in unaffected siblings and family members of ASD patients [200]. Considering ASD a continuum trait also in the general population will enable the study of intermediate levels of ASD in larger more accessible samples and the investigation of the underlying mechanisms of the disorder. Actually, this approach has been beneficial for depression research: studying subclinical levels of depression highlights common genetic mechanisms [201]. As not many biological cohorts have included a continuum measure of ASD, a possible solution would be to analyse the genetic potential for ASD in the general population using polygenic risk scores (PRS). PRS uses the sum of all known common variants that contribute to a disorder to calculate an overall risk of getting this disorder. As mentioned before, recent research has shown that the polygenic form of autism is made up of the additive effects of individual SNPs. Therefore, using this methodology, the mechanisms of autistic-like traits could be studied in any cohort with genetic and MRI imaging information available. To date, several studies have revealed that PRS for ASD correlate with cognitive abilities, such as logical memory or executive function [202,203], and brain structural alterations (cortical thickness and white matter connectivity) [204].

## 2.4.  Syndromic autism

The traditional definition of syndromic autism is a disorder with a clinically defined pattern of somatic abnormalities and a neurobehavioural phenotype that may include ASD. Approximately 5-10% of ASD patients have co-occurring monogenic syndromes or disorders and the diagnosis is typically confirmed by targeted genetic testing. Syndromic autism may be caused by loss-of-function of specific genes or chromosomal abnormalities, examples are reported in Box 2 and Figure 13.

**BOX 2. Genetics of syndromic autism**

**Fragile X syndrome** (OMIM #300624) is caused by the repetition (more than 200 times) of the trinucleotide CGG at 5' of the *FMR1* gene (Xq27.3). These repetitions lead to hypermetylation of the region and silencing of the *FMR1* gene. Patients present characteristic dysmorphic features accompanied with neurological problems and, in many cases, with ID and/or autism.

*MECP2* encodes MeCP2 (methyl-CpG-binding protein 2), a protein that binds to methylated CpG islands to suppress transcription. Loss-of-function mutations in *MECP2* cause 70% of **Rett syndrome** (OMIM #312750) in girls, these mutations being lethal in boys. Depending on the mutation, its location in the gene and the inactivation pattern of X chromosome in each individual, the phenotypic traits may differ: ID, ASD, learning problems or even asymptomatic presentations.



**Figure 13. Prevalence of ASD in syndromes of known genetic aetiology**. de la Torre-Ubieta et al., 2016.

**Tuberous sclerosis complex**, TSC, (OMIM #191100) is an autosomal disorder characterized by the presence of benign tumours in the brain and other organs. The disorder is caused by mutations in *TSC1*, encoding hamartin, or *TSC2*, encoding tuberin. Patients with TSC present a wide range of neurodevelopmental disorders such as ASD, present in 50% of cases.

**Neurofibromatosis type I** (OMIM #162200) is an autosomal dominant disorder caused by mutations in *NF1*, situated in the 17q11.2 region. This gene encodes neurofibromin, a tumour suppressor protein. Patients with mutations in *NF1* have a higher probability to develop benign and malign tumours in the nervous system. Mutations in *NF1* lead to the activation of the mTOR signalling pathway and neuronal hypertrophy, which may cause impaired neuronal connectivity and synaptic plasticity, associated with ASD.

**Adenylosuccinate lyase deficiency** (OMIM #103050) is a metabolic syndrome associated with ASD. This autosomal recessive disorder is caused by mutations in *ADSL*, encoding adenylosuccinate lyase. This syndrome has a heterogeneous phenotype with psychomotor impairment, epileptic convulsions and ASD in 80-100% of cases.

Patients with syndromes caused by **chromosomal abnormalities** are often diagnosed with ASD. Structural chromosomal alterations have been reported for every chromosome and include deletions, duplications, inversions or translocations. The most frequent chromosomal abnormality detected in 1–3% children with ASD is a maternally inherited **15q11-q13** duplication. Many genes in this chromosomal

region have essential functions in the brain, such as *GABRA5* and *GABRB3* (GABA receptors), *UBE3A* and *HERC2* (components of the proteasome complex) and *SNRPN* (ribonucleoprotein peptide N) as well as *CYFIP1* (the FMRP interacting protein)

## 2.5.    Idiopathic autism: candidate genes

With the development of sequencing technologies, over hundreds of candidate causal genes of ASD have been identified. The SFARI database currently reports more than 1,000 genes associated with ASD (around 5% of the protein-coding genes in the genome), 207 showing high-confidence associations with ASD. Interestingly, in this database candidate genes are classified into different categories, annotated for their relevance to autism and scored according to the strength of the evidence for each gene's association with ASD (Table 4).

**Table 4. Top 10 genes reported to be associated with ASD in SFARI with high confidence (score 1).**

| Gene Symbol | Chromosome | ASD reports/ Total reports | Rare variants/ Common variants |
|:---:|:---:|:---:|:---:|
| *NRXN1* | 2p16.3 | 47/90 | 212/4 |
| *SHANK3* | 22q13.33 | 47/89 | 261/9 |
| *MECP2* | Xq28 | 27/87 | 180/0 |
| *SCN2A* | 2q24.3 | 40/72 | 290/0 |
| *SCN1A* | 2q24.3 | 25/68 | 158/2 |
| *PTEN* | 10q23.31 | 26/64 | 129/0 |
| *CHD8* | 14q11.2 | 40/61 | 215/0 |
| *SYNGAP1* | 6p21.32 | 21/60 | 172/0 |
| *GRIN2B* | 12p13.1 | 25/57 | 154/32 |
| *RELN* | 7q22.1 | 29/54 | 166/9 |

The 10 genes with score 1 reported in a higher number of studies are listed.

Also, the largest exome sequencing study to date, with a combined sample of more than 35,000 subjects, has brought the number of high-confidence genes up to 102 [205] (Figure 14). Despite their diversity, many of these ASD risk genes are highly enriched in pathways related to important cellular functions - chromatin remodelling, transcription and alternative splicing - and neuronal function – neurogenesis, synaptic plasticity, neuronal connectivity, and migration [115,182,206]. (Figure 15)
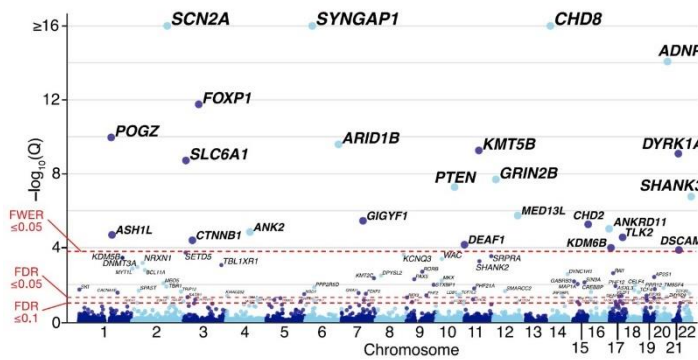


**Figure 14. 102 autosomal genes associated with ASD in Satterstrom et al., 2020.** In this Manhattan plot, each point represents a gene. 78 of these genes overcome false discovery rate (FDR)≤0.05 and 26 overcome family-wise error rate (FWER)≤0.05. Adapted from Satterstrom et al., 2020 [205].
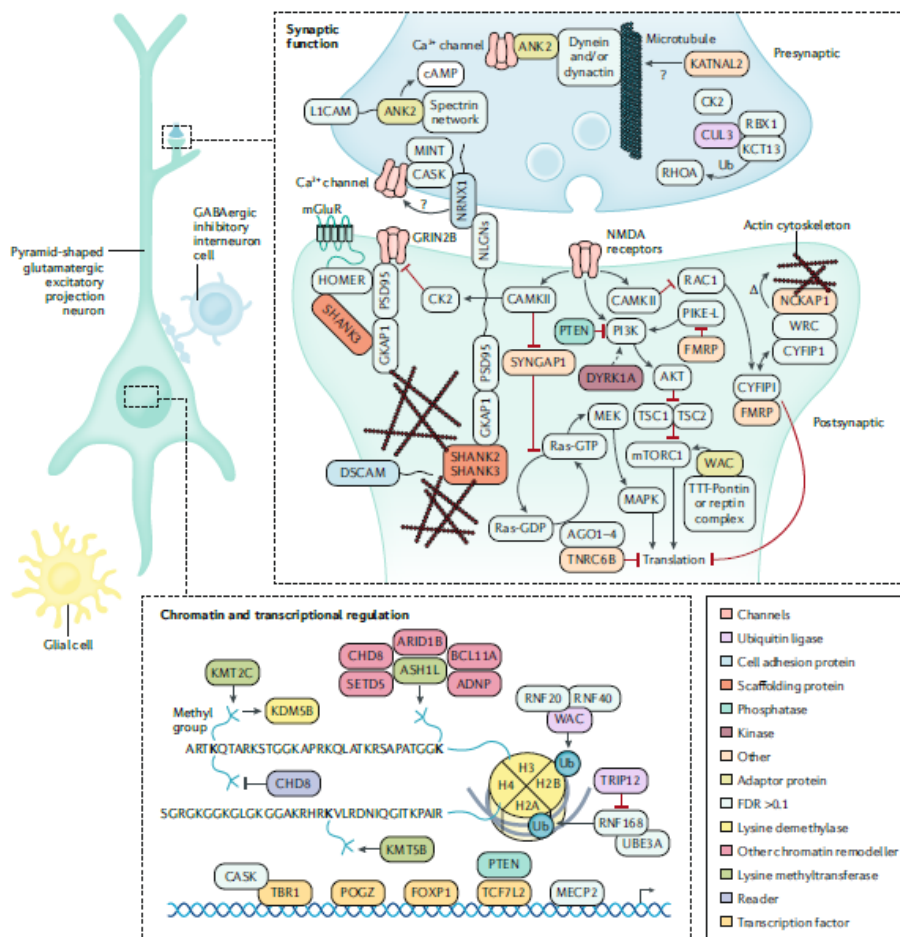


**Figure 15. Encoded proteins and pathways associated with ASD.** Major components of a neuronal circuit in the cerebral cortex, with a focus on pyramidal glutamatergic neurons. Proteins encoded by selected high-confidence ASD-risk genes are represented as false discovery rate (FDR) >0.1. Lord, 2020 [22].

In this thesis we have focused on several gene families (14-3-3 and *BEX*/*TCEAL*) and candidate genes (*RBFOX1*) to investigate their contribution to ASD and co-occurrent psychiatric disorders.

**The 14-3-3 gene family**

14-3-3 is a protein family of seven members (β, γ, ε, η, ζ, σ, θ) that act as molecular chaperones and interact with hundreds of other proteins. 14-3-3 proteins are involved in important biological processes such as cell cycle, transcription, neuronal development, migration and neurite outgrowth [207–209]. Some members of this family seem to play an important role in neurogenesis and neurodifferentiation, and may be implicated in neurodevelopmental disorders [208,210]. Interestingly, microduplications of *YWHAE*, which codes for 14-3-3 ε, have been reported in ASD patients and microdeletions involving *YWHAE* cause Miller-Dieker syndrome, a form of lissencephaly. In addition, recent studies have associated several 14-3-3 genes to psychiatric disorders, such as ADHD, bipolar disorder, major depressive disorder, schizophrenia and suicide attempts [211–219]. Finally, knock-out mouse models of 14-3-3 genes have been studied, showing that animals deficient for 14-3-3 proteins present a variety of behavioural manifestations related to psychiatric and neurodevelopmental disorders [211,220–226] (Table 5).

**Table 5. Summary of phenotypical and neurological alterations found in knockout mouse models of different 14-3-3 genes.**

| Knockout isoform | Behavioural phenotype | Neurological alterations | Refs. |
|---|---|---|---|
| **14-3-3 γ** | Hyperactivity, depressive-like behaviour, more sensitive responses to acute stress. | Developmental delay | [224] |
| **14-3-3 ε** | Defects in working memory, moderately high anxiety-like behaviour, increased locomotor activity, increased sociability | Defects in brain development and neuronal migration | [211,221,223] |
| **14-3-3 ζ** | Impaired learning and memory, locomotor hyperactivity and disrupted sensorimotor gaiting | Developmental abnormalities in hippocampus. Aberrant neuronal migration. Abnormal mossy fibre navigation and glutamatergic synapse formation. Enlarged lateral ventricles and reduced synaptic density. | [220,222] |
| **14-3-3 ε and ζ** | Severe phenotypes during embryogenesis. Seizures. | Increased number and aberrant distribution of progenitor cells in the developing cerebral cortex. Increased differentiation of neural progenitor cells into neurons. Neuronal migration defects in cortex. | [226] |

An exome sequencing study performed by our group in ASD multiplex families reported a series of gene-disrupting rare variants shared by the affected siblings in each family, one of them found in *YWHAZ*. Interestingly, a protein-protein interaction analysis including all genes found mutated in this study reported YWHAZ as the main node of the network (Figure 16), which has prompted us to investigate its contribution to ASD and other comorbid disorders in the context of this thesis.
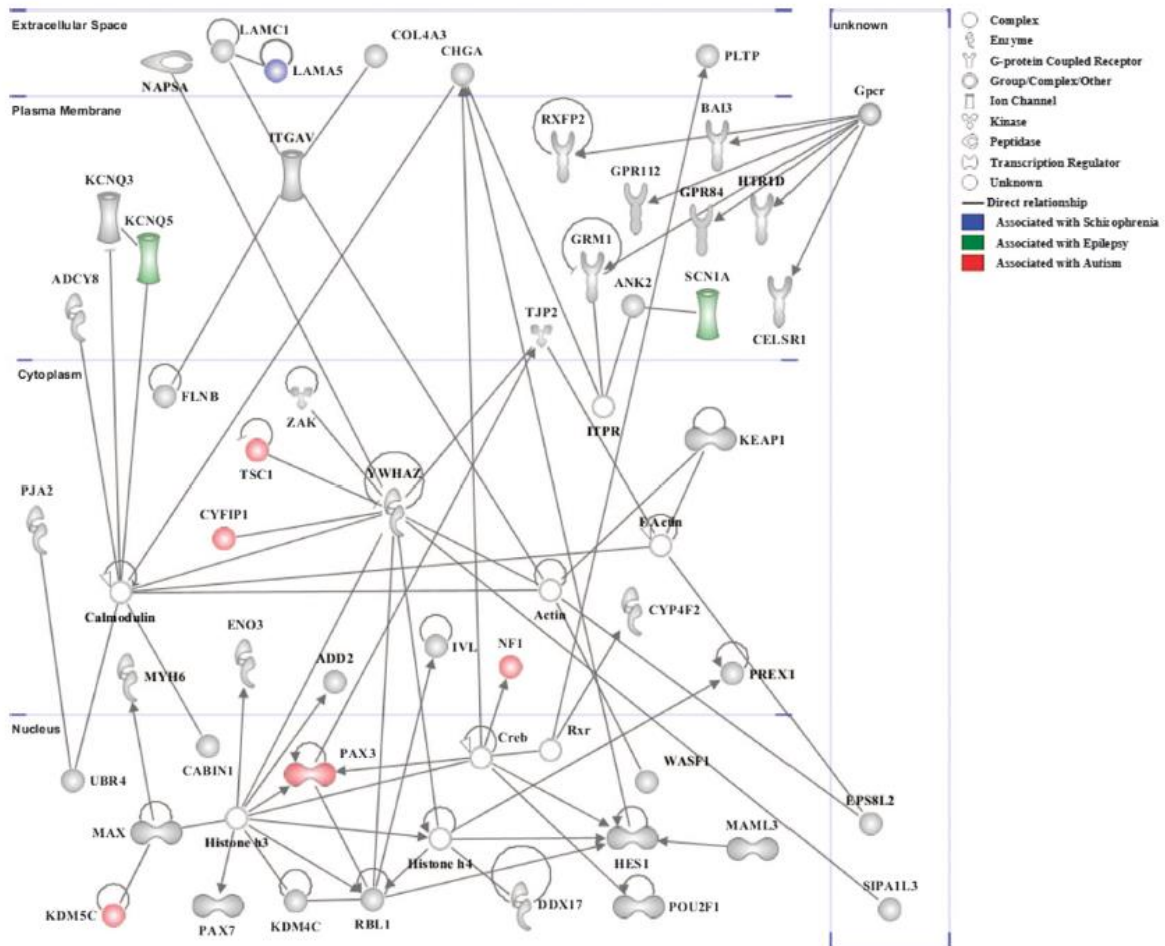


**Figure 16. Protein-protein interaction analysis performed with all the mutated genes found in Toma et al., 2014** [227] **highlights YWHAZ as the main node**. Toma et al., 2014 [227]

### *RBFOX1*

*RBFOX1* (also referred to as *A2BP1* or *FOX1*) encodes a RNA splicing factor, specifically expressed in brain, heart and muscle (GTEX, www.gtexportal.com), that regulates the expression and splicing of large gene networks during early neuronal development [228–230]. RBFOX1 has two main isoforms, a cytoplasmic and a nuclear one, that seem to play different roles.

- The nuclear isoform would contribute to mRNA stability and promote translation. It is involved in neuronal migration and synapse network formation during corticogenesis and is important for the control of neuronal excitation in the mammalian brain [231,232].
- The cytoplasmic isoform acts as a splicing regulator of a large number of genes [233].

Alterations in the *RBFOX1* gene have been associated with several neurodevelopmental pathologies, especially with ASD [228,234]. Several studies have reported CNVs and point damaging mutations in *RBFOX1* in autistic patients [163,235–237], and transcriptomic analysis of *postmortem* autistic brains have revealed decreased levels of *RBFOX1* and dysregulation of *RBFOX1*-dependent alternative splicing [238]. Also, *RBFOX1* haploinsufficiency results in a syndrome characterized by neurodevelopmental and neurological phenotypes including ASD, ID and epilepsy [239,240].

This gene seems to be implicated not only in autism, but also in other psychiatric traits. For instance, genetic and neuroimaging studies, together with animal models, have related *RBFOX1* to aggressive phenotypes [241]. Furthermore, common and rare genetic variants in *RBFOX1* have been associated in multiple reports to other psychiatric conditions, such as major depression, schizophrenia or ADHD [242–244]. Interestingly, *RBFOX1* emerges as a top hit in a GWAS meta-analysis of eight psychiatric disorders [192].

In addition, data from cellular and animal models highlight the involvement of *RBFOX1* in neurodevelopment and synaptic transmission. First, *RBFOX1* knockdown of human neural progenitor cells modelling haploinsufficiency during neuronal differentiation demonstrated widespread changes in RNA splicing and gene expression [245]. Also, knockdown of Rbfox1 proteins in mouse neurons and subsequent rescue with cytoplasmic or nuclear isoforms demonstrated that cytoplasmic *Rbfox1* isoform regulates the expression of synaptic and autism-related genes [233]. Finally, *Rbfox1* neural-specific KO mouse shows alterations in synaptic transmission, increased excitability, and a predisposition to seizures [232], and specific knockdown of the nuclear *Rbfox1* isoform in mouse neurons demonstrated the critical role of this isoform in neuronal migration and synapse network formation during corticogenesis [231].

**The *BEX/TCEAL* gene family**

The *BEX/TCEAL* gene family consists on a genetic cluster, spanning around 1.5Mb in the Xq22 locus, formed by 5 *BEX* (brain-expressed X-linked) and 9 *TCEAL* (transcription elongation factor A (SII)-like) genes. All *BEX* genes and some *TCEAL* genes are highly expressed in brain tissues

(GTEX, www.gtexportal.org) and, although not well studied, BEX proteins seem to play a role in neuronal development.

BEX proteins are suggested to be involved in neurotrophin receptor signalling, neurotrophins being important neuronal growth factors involved in the development, maintenance, survival, differentiation and apoptosis of the nervous system [246]. Also, an elevated *BEX1* expression was detected in spinal motor neurons (MNs) after axonal damage and in mutant mice with MNs degeneration [247,248], and *BEX1* knockout mice show a lower recovery from sciatic nerve injury [248], data that suggest *BEX1* involvement in neuronal regeneration.

Finally, several reports point to a role of the *BEX* gene family in several neurological conditions. In a recent analysis, *BEX1* and *BEX3* have been identified among the most significantly downregulated genes in excitatory neurons of the prefrontal cortex of patients with Alzheimer's disease [249]. Furthermore, microdeletions and microduplications including *BEX/TCEAL* genes have been described in autistic patients. Microdeletions in Xq22, encompassing *BEX/TCEAL* genes, were found in patients with neurodevelopmental problems, severe ID or autism, with BEX3 highlighted as one of the main candidates to cause these neurological traits [250–252], and a small 252-kb duplication spanning *BEX3*, *TCEAL4*, *TCEAL9* and *RAB40A* was reported in a patient with ASD in the Decipher database (DECIPHER ID: 290829, www.dechipergenomics.org).

# 3. ANIMAL MODELS OF ASD

Animal models are a helpful tool to investigate genetic and neuronal mechanisms of human psychiatric and neurodevelopmental disorders, as they allow to explore behavioural and cognitive alterations in a complex organism. Rodents are the most used animal model in neuroscience research, although zebrafish models are becoming popular as they present several advantages [253]. To study the neurobiology of ASD in animal models two main approaches can be followed: forward genetics, in which ASD-like phenotypes are identified in the animal and then the molecular basis of these alterations are investigated, or reverse genetics, the most commonly used, in which targeted mutations are introduced into the genome of the animal and then its phenotype is characterized.

## 3.1.   Rodent models of ASD

Rodents, especially mice, have been widely used as models of human disorders due to the early sequencing of their genome, back in 2001, and the fact that they can be easily maintained in an animal facility. The protein-coding regions of the mouse and human genomes are 85% identical, figure that is similar in rats. Moreover, the Mouse Genomes Project from the Wellcome Sanger Institute (www.sanger.ac.uk/data/mouse-genomes-project/) has sequenced the genomes of the most highly used laboratory mouse strains in order to facilitate genetic research.

Mice have a well-described development and life-cycle divided into: prenatal stages (-19 days to date of birth), early postnatal development (0-21 postnatal days, PND), adolescence (21-60 PND) and adulthood [254]. Importantly, they are social animals and present complex behavioural phenotypes that can be assimilated to human behaviours. Indeed, tests in rodents enable to identify behaviours comparable to the core symptoms of ASD: impairments in social interaction, communication and presence of repetitive behaviours [255–257] (Table 6). Also, recent advances in molecular genetics have permitted the easy manipulation of the rodent genome, making it possible to study the effects of loss of function of specific genes in behaviour and neurobiology. Rodent genetic models stand as a fundamental preclinical tool to clarify the complex aetiology of ASD and to test new potential treatments before clinical trials [256,258].

Some mouse strains with ASD-like phenotypes have arisen due to inbreeding procedures, such as the BTBR T+tf/J strain [259]. However, the majority of ASD mouse models used to date have been generated by altering specific ASD-linked genes in the mouse genome. There are nearly 200 mouse models developed to study candidate genes, some of them recently reviewed in published articles and in the SFARI database (https://gene.sfari.org/) [253,255,260,261].

**Table 6. Behavioural assays to evaluate ASD-like alterations in rodents and zebrafish**. Adapted from Pensado-López, 2020 [253].

| Areas of interest | Behavioural assays in rodents | Behavioural assays in zebrafish |
|---|---|---|
| **Socialization** | Social approach task: time spent with an unknown individual compared to a new non-social object | Social preference and social novelty tests: preference for interacting with conspecific strangers, and preference for interacting with familiar conspecifics or with unfamiliar ones |
| | Social preference and social novelty tests (affiliation and recognition): time spent with an | Shoaling test: measure of the swimming behaviour of a group of individuals (nearest |

| | | |
|---|---|---|
| | unknown animal in comparison with a familiar one | neighbour distance, interindividual distance, cluster score) |
| | <u>Free interaction test</u>: time spent interacting with unknown individuals compared to the time spent doing other activities (exploring…) | <u>Social interactions</u>: presence of behaviours such as approaching, circling, mouth opening, biting, chasing, etc. |
| | <u>Partition test</u>: measure of the time spent close to the stranger individual separated by a partition | |
| | <u>Reciprocal social interactions</u>: presence of interactions such as sniffing, following, pushing each other, etc. | |
| **Non-social patterns of behaviour** | <u>Open-field test</u>: presence and duration of repetitive patterns of locomotor activity | |
| | <u>Reversal learning tasks (insistence for sameness)</u>: measure the flexibility of the mouse to switch from an established habit to a new habit, using T or Y-maze, or Morris Water Maze | <u>Open-field test</u>: presence and duration of repetitive patterns of locomotor activity |
| | <u>Restricted interests</u>: measure of exploratory activity of the animal in a novel environment or exposed to novel objects | <u>Inhibitory avoidance response</u>: measure of the latency of an individual to enter a chamber with an aversive response |
| | <u>Burying behaviour</u>: presence of digging behaviour | |
| | <u>Repetitive self-grooming</u> | |
| **Communication** | <u>Ultrasonic vocalizations (USVs)</u>: reduced levels of USVs or non-usual patterns of acoustic communication, or altered response to them | Non-available |
| | <u>Habituation and dishabituation to social odours</u>: response to a change in a familiar odour for a new one | |

## 3.2. Zebrafish as a model of ASD

The zebrafish was initially used to study vertebrate development but, in recent years, it has been established as a powerful model for studying neurological and psychiatric diseases [262–266]. Zebrafish present some advantages when compared to rodent models. First, they are highly prolific and develop rapidly, which enables high-throughput assays. The zebrafish lifecycle is divided into four main periods: embryo (0-3 days post fertilization, dpf), larva (3 dpf – 6 weeks post fertilization), juvenile (6 weeks – 3 months post fertilization) and adult. Second, the external fertilization and transparency of embryos and larvae allow to study neurodevelopmental processes in the intact brain using optogenetics or imaging techniques. Finally, its maintenance is relatively easy and cost-effective [266].

An important consideration when working with zebrafish is that during evolution teleost underwent an additional whole genome duplication and, therefore, many human genes present more than one orthologue in the zebrafish genome. However, zebrafish stands as a good genetic model due to its high genetic similarity to human, and around 70% of human genes have at least one zebrafish orthologue [267]. Importantly, the Zebrafish Genome Project from the Wellcome Sanger Institute (https://www.sanger.ac.uk/data/zebrafish-genome-project/) have created reference assemblies for zebrafish strains and therefore facilitate the use of this animal model in genetic research.

Neurodevelopmental events in zebrafish are also similar to humans and, although timings and brain organization differ, comparative studies have precisely mapped these differences, allowing to transfer the information gained in zebrafish to other species [268–270]. Importantly, zebrafish display a well-defined social behaviour, established early during development and maintained throughout life. They express strong preferences towards conspecifics and live in mixed-gender groups with structured dominance hierarchies, which makes zebrafish an interesting model of ASD. A battery of tests for behavioural analysis of both larval and adult zebrafish have been developed in recent years, allowing to phenotype genetic models of human disorders such as ASD [269,271] (Table 6).

Another interesting advantage of zebrafish is the relative ease and versatility to conduct genetic manipulations in embryos. The most commonly used genetic approaches in zebrafish are morpholino-based expression silencing, ENU-based mutagenesis, and, more recently, CRISPR/Cas9 gene editing [269,272]. Using these genetic manipulation techniques, an increasing number of zebrafish models of ASD has been created in the last years. Importantly, the Zebrafish Mutation Project (zmp.buschlab.org) from the Wellcome Sanger Institute created a repository

of over 40,000 mutant zebrafish lines with mutations covering 60% of zebrafish protein-coding genes that are available to order and use for research purposes. Some of these models exhibit neurological and behavioural alterations and are reviewed in several published papers [269,272,273].

# 4. BRAIN ACTIVITY AND CONNECTIVITY IN ASD

Brain imaging approaches, such as magnetic resonance imaging (MRI), positron emission tomography (PET) or single-positron emission computed tomography (SPECT), together with electrophysiological methods like electroencephalography (EEG) and magnetoencephalography (MEG), offer excellent non-invasive ways to investigate the effects of genetic variation on brain structure, function, and connectivity directly in humans *in vivo* (Box 3). The complementary characteristics of these techniques enable gathering information on how genetic alterations affect brain anatomy and function, in terms of location (especially MRI), neurochemical changes (PET and SPECT) and timing (especially EEG and MEG). Imaging and electroencephalographic studies in patients with neurological or psychiatric disorders have complemented genetic and molecular findings and contributed to depict the complex neurobiology underlying these disorders

---

**BOX 3. Brain imaging and electrophysiological techniques**

**MRI** is a versatile imaging technique based on the special spin properties of protons and neutrons.

- **Structural magnetic resonance imaging (sMRI)** allows to noninvasively characterize the structure of the human brain with high spatial resolution. This technique uses the different magnetic properties of brain tissues to map *in vivo* the spatial distribution of brain areas and subtle changes in brain morphology.
- **Diffusion tensor imaging (DTI)** uses free water diffusion to investigate the axonal white matter organization of the brain, and **arterial spin labelling (ASL)** quantitatively measures tissue perfusion.
- **Functional magnetic resonance imaging (fMRI)** can be used to investigate potential changes in brain activity **after a task-induced stimulus or during a resting condition (rs-fMRI)**. fMRI measures a blood-oxygen level-dependent (BOLD) signal which is thought to be proportional to brain function levels, although the relationship between cell activation, oxygen saturation, and cerebral blood flow changes is arguable [274]. fMRI is useful to analyse temporal

---

correlations of neural activity across distinct brain regions and functional connectivity based on spontaneous or induced brain activity, neural organization, and circuit architecture.

**PET and SPECT** are molecular imaging techniques that use radiolabelled tracers to study molecular interactions of biological processes *in vivo*. PET is more commonly used due to its higher sensitivity and temporal resolution compared to SPECT. These techniques can be used to assess glucose metabolism in specific areas of the brain, as a correlate of brain activity, to measure neurotransmitter levels, or to follow specific molecular processes with radiotracers.

**EEG and MEG** are electrophysiological techniques that record electrical activity of the brain with high temporal resolution, which allows to study brain dynamics at millisecond timescales and to establish direct relationships with neuronal activity (which is not the case in fMRI, that measure blood dynamics). These techniques provide valuable information about brain connectivity, dynamic activation and deactivation of functional networks. EEG is more economic, easier to use and less invasive than MRI. MEG is more expensive but provides higher spatial resolution than EEG.

## 4.1. Findings from MRI studies

In the last years, structural and functional brain imaging studies have accelerated our understanding of the relationship between altered neural circuits and clinical symptoms in autistic patients [275].

On one side, sMRI is a useful technique to study morphological alterations of the whole brain or specific areas during neurodevelopment. Indeed, sMRI studies on ASD have detected differences in brain growth patterns, grey matter and white matter volumes in young children with autism, particularly in the frontal cortex, temporal cortex and amygdala, and suggest the presence of disrupted neural pathways before the appearance of behavioural symptoms in these children [38,39,276–282].

On the other side, fMRI allows to investigate the functional activity of specific areas suspected to be involved in autism. Task-based fMRI studies have found differences in activation between autistic patients and controls in areas related to language processing and integration of sensory information [283–285]. In addition, rs-fMRI studies are used to study intrinsic connections in the human brain. In these studies, participants look at a blank screen with no task demands and brain activity is recorded in order to find differences in basal activity and connectivity between patients and controls. To allow highly powered studies in ASD, large data sets have been pooled, such as the Autism Brain Imaging Date Exchange (ABIDE,

fcon_1000.projects.nitrc.org/indi/abide) [286]. Interestingly, these resting-state studies have found evidence of both hyper-connectivity and hypo-connectivity in short-range connections throughout the brain such as the salience network, default mode network, executive control network, dorsal attention network, or the corticostriatal and vasopressin-related circuits [275,287,288]. Finally, ASL studies have detected altered resting functional connectivity and cerebral perfusion in both grey and white matter in ASD patients compared to control children [275,289,290].

All these evidence from MRI studies on structural and functional brain alterations would reflect neurodevelopmental problems and therefore contribute to the later symptomatology observed in patients (Figure 17). However, there is heterogeneity in the results obtained, and the majority of the studies are limited by averaging data across many individuals (which can mask heterogeneity and differences across age groups), small sample sizes and problems with replication.
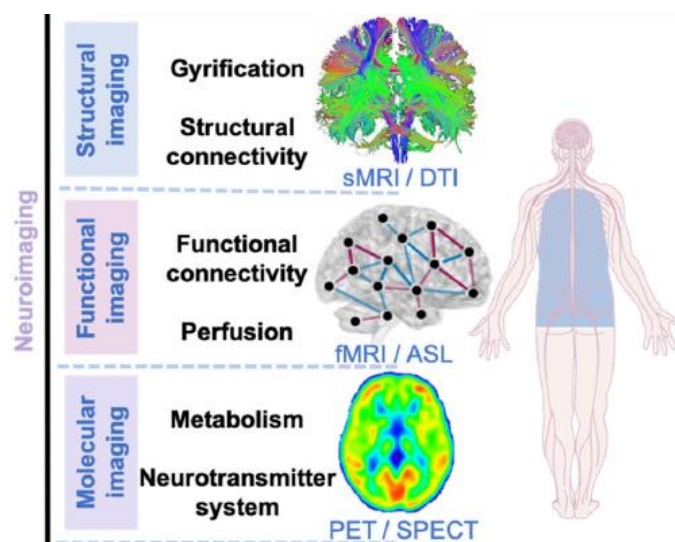


**Figure 17. Neuroimaging studies on ASD patients help to follow brain development abnormalities**. The different neuroimaging techniques are listed together with the dimensions they assess. Adapted from Li et al., 2021 [275].

## 4.2. Findings from molecular imaging studies

Several studies have used PET to analyse task-dependent glucose metabolism in ASD patients, as a correlate of brain activity. These studies found a decreased metabolism in the amygdala, frontal premotor and eye-field areas, and parietal lobe in ASD patients compared to controls, which could be related to hypoactivation of these areas during a cognitive task [291]; and increased glucose metabolism in several white matter structures of autistic patients [292]. This increased glucose metabolism might reflect inefficient functioning of these areas. In addition, PET and SPECT studies in ASD patients have reported abnormalities in neurotransmitter levels,

mainly in 5-HT, DA, GABA and glutatamate systems, and in neuroinflammation. All these findings are reviewed in Li, 2021 [275] (Figure 17).

## 4.3.   Findings from electroencephalographic studies

EEG has historically been used for the diagnosis of comorbid epilepsy in people with autism. Later, some task-based EEG studies were performed to study the neurological mechanisms of autism itself. Initially, they focused on the modulation of cognitive function that relates to the autistic phenotype, following the "broken mirror" theory of autism. This theory postulates that autistic patients are unable to "mirror" observed actions, and is based on altered mu wave suppression [293]. However, subsequent studies questioned this theory [294–297], pointing to a more complex picture of dysfunctional executive functions and visual attention [298]. Other task-based studies demonstrated a differential sensory processing in autistic individuals, observing changes in sensitivity and latency and suggesting that differences in auditory and visual processing could be related to language delay and difficulty in emotion recognition in patients [299,300].

Interestingly, a recent review in EEG and MEG studies reports several major findings regarding ASD brain connectivity. First, ASD is characterized by a more randomly organized functional connectivity. Second, an abnormal lateralization of functional connectivity has been reported in patients, especially an increased left-over-right EEG and MEG connectivity ratio. In neurotypical subjects, some cognitive tasks involve the specific activation of one cerebral hemisphere and not the other. In ASD patients, alterations in brain circuitry may impair this specialized lateralization and affect cognition and behaviour. Third, both abnormal inter- and intra-hemispheric connectivity has been described in ASD patients: mainly a long-range underconnectivity that may be compensated by a short-range overconnectivity [77]. Finally, several studies have demonstrated that differences in EEG in high-risk infants may represent endophenotypes of autism [301,302].

## 4.4.   Brain imaging in zebrafish models

In the last years, chemical dyes and genetically encoded fluorescent proteins have been developed to study neuronal activity at a single-cell resolution in animal and cellular models. Among them, GCaMP calcium indicators have become the most widely used [303,304]. The

combination of these sensitive indicators with a two-photon laser scanning fluorescence microscope, which can penetrate deep into scattering tissues, and light sheet or light field microscopy (LFM), that allows to cover a large imaging volume at high speed, has enable to record neuronal activity in animal and cellular models *in vivo* [305]. These imaging techniques have a high spatial resolution and are non-invasive, which makes them powerful tools to investigate alterations in neuronal activity and connectivity in different animal models of human disorders.

As mentioned before, the small size and transparency of zebrafish larvae, and the possibility to use mutant transparent lines, allow to record the activity of the whole brain *in vivo* using neuroimaging approaches. In the last decade, the improvement of imaging techniques has enabled to study whole-brain activity in zebrafish larvae in resting states, after exposure to simple *stimuli*, or even in freely moving individuals [306–311] (Figure 18). Whole-brain functional imaging with cellular resolution is a powerful tool to study neurological alterations in genetic zebrafish models of psychiatric and neurodevelopmental diseases. To date, *in vivo* whole-brain imaging studies have been focused on studying the neural circuits implicated in the response to *stimuli* or in specific motor responses, and only few studies have investigated differences between wild-type and mutant fish for specific genes [312,313]. However, this new technique stands as an interesting tool to explore the effects of mutations in specific genes in neuronal activity and connectivity and, therefore, to contribute to our understanding of the neurobiology of disorders such as ASD.
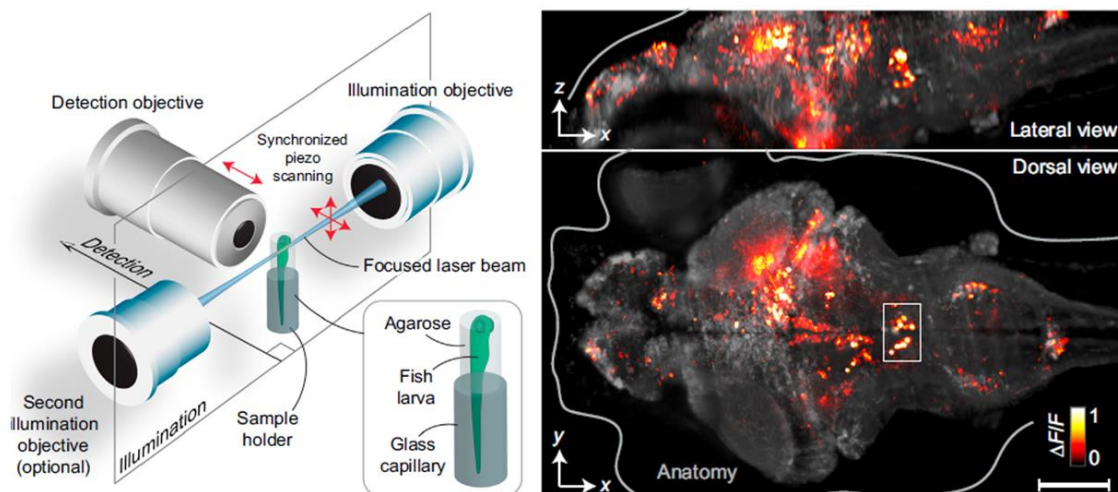


**Figure 18. Whole-brain imaging in larval zebrafish.** On the left, an example of a light-sheet imaging setup for whole-brain imaging recordings in larval zebrafish. On the right, a 3D reconstruction of the whole zebrafish brain activity during a neuroimaging recording using a GCaMP5G calcium marker. Changes in fluorescence intensity levels are indicated in yellow and red. Keller et al., 2015 [309].

# OBJECTIVES

The overall objective of this work is to investigate the contribution of candidate genes to autism spectrum disorder (ASD) and comorbidities and to gain insight into the genetic and neurological basis of these disorders. We aim to determine the contribution of common and rare variants in the 14-3-3 gene family, *RBFOX1* and the *BEX/TCEAL* gene family to ASD and other psychiatric disorders and to functionally characterize the effect of their deficiency using animal models.

The specific aims of this work are:

**CHAPTER 1**. Exploring the contribution of the 14-3-3 gene family to ASD and other psychiatric disorders

1.1. Evaluate the contribution to ASD and other psychiatric disorders of common and rare variants in the 14-3-3 gene family

1.2. Explore possible altered expression levels of the 14-3-3 gene family in the brain of *postmortem* ASD and schizophrenia patients

1.3. Evaluate the functional effect of a specific mutation in the *YWHAZ* gene reported in two siblings diagnosed with ASD and attention deficit/hyperactivity disorder (ADHD) using *in vitro* assays

1.4. Investigate *ywhaz* expression across development and in adulthood in zebrafish

1.5. Assess the effect of loss of *ywhaz* function on neural activity and connectivity, neurotransmission and behaviour using a zebrafish knockout (KO) model

**CHAPTER 2**. Exploring the contribution of *RBFOX1* to ASD and other psychiatric disorders

2.1. Evaluate the contribution to ASD and other psychiatric disorders of common variants and rare copy number variants in *RBFOX1*

2.2. Explore altered expression levels of *RBFOX1* in the brain of *postmortem* ASD and schizophrenia patients

2.3. Investigate the effect of the rs6500744 variant, situated in an intronic region of *RBFOX1*, in brain circuitry using fMRI

2.4. Evaluate the effect of loss of *Rbfox1* function on behaviour using a mouse KO model

2.5. Investigate *rbfox1* expression across development and in adulthood in zebrafish

2.6. Assess the effect of loss of *rbfox1* function on neural activity, connectivity, and behaviour using two zebrafish KO models

**CHAPTER 3**. Exploring the contribution of the *BEX/TCEAL* gene family to ASD and other psychiatric disorders

3.1. Assess the contribution of common and rare variants in *BEX3* to ASD and schizophrenia using public data and a sequencing a cohort of ASD patients

3.2. Explore altered expression levels of the *BEX/TCEAL* gene family in the brain of *postmortem* ASD and schizophrenia patients

3.3. Assess the effect of loss of *Bex3* function on brain morphology and function and on behaviour using two KO mouse models

# RESULTS

# THESIS SUPERVISORS' REPORT ON THE CONTRIBUTION OF THE PHD CANDIDATE TO THE ARTICLES INCLUDED IN THIS DOCTORAL THESIS

**Thesis title**: Pleiotropic effects of candidate genes on autism spectrum disorder and comorbidities: genetics, functional studies and animal models

**Author**: Ester Antón Galindo

**Supervisors**: Noèlia Fernàndez Castillo and Bru Cormand Rifà

The undersigned, Noèlia Fernàndez Castillo and Bru Cormand Rifà, acting as PhD supervisors of the Thesis by Ester Antón Galindo entitled: "Pleiotropic effects of candidate genes on autism spectrum disorder and comorbidities: genetics, functional studies and animal models" and presented as a compendium of 5 research articles, hereby inform about the precise contribution of the candidate in each publication.

**CHAPTER 1: Exploring the contribution of the 14-3-3 gene family to ASD and other psychiatric disorders**

Article 1:

**Title**: Involvement of the 14-3-3 gene family in autism spectrum disorder and schizophrenia: genetics, transcriptomics and functional analyses

**Authors**: Torrico B*, Antón-Galindo E*, Fernàndez-Castillo N*, Rojo-Francàs E, Ghorbani S, Pineda-Cirera L, Hervás A, Rueda I, Moreno E, Fullerton JM, Casadó V, Buitelaar JK, Rommelse N, Franke B, Reif A, Chiocchetti AG, Freitag C, Kleppe R, Haavik J, Toma C[#], Cormand B[#].

* equally contributed to this work; [#] equally supervised this work

**Journal**: Journal of Clinical Medicine. 2020 Jun 13;9(6):1851. doi: 10.3390/jcm9061851.

**Impact Factor (2019)**: 3.303

**Contribution of the PhD candidate**:

Cloning and characterization of plasmids for the BRET assays. Setting up the BRET protocol in the laboratory and performing the BRET assays to test 14-3-3ζ and 14-3-3ε heterodimerization, 14-3-3σ homodimerization and 14-3-3σ interaction with DYRK1A. Analysis of GWAS data. Exhaustive collection of transcriptomic data from publications and GEO datasets. Writing the first draft of the manuscript, figures and tables, and participating in the final edits.

**Article 2**:

**Title**: Deficiency of the *ywhaz* gene, involved in neurodevelopmental disorders, alters brain activity and behaviour in zebrafish

**Authors**: Antón-Galindo E*, Dalla Vecchia E*, Orlandi JG, Castro G, Loza-Alvarez P, Aguado F, Norton WHJ[#], Cormand B[#], Fernàndez-Castillo N[#]

* equally contributed to this work; [#] equally supervised this work

**Journal**: To be submitted to Molecular Psychiatry

**Contribution of the PhD candidate**:

Crossing and genotyping zebrafish to obtain the transgenic lines for the brain imaging analysis. Quantification of GCaMP6s expression by qPCR. Setting up the protocols for the whole-brain imaging recordings and analysis. Setting up and performing the analyses of whole-brain imaging recordings, extraction of fluorescence signals, definition of brain areas and statistical analysis of brain activity and connectivity. Performing two out of three replica of the behavioural tests. Writing the first draft of the manuscript, figures and tables, and participating in the final edits.

**CHAPTER 2: Exploring the contribution of *RBFOX1* to ASD and other psychiatric disorders**

**Article 3**:

**Title**: One gene to rule them all: *RBFOX1* and mental disorders

**Authors**: O'Leary A, Fernàndez-Castillo N, Gan G, Antón-Galindo E, Cabana-Domínguez J, Yotova A, Kranz T, Grünewald L, Burguera D, Pané-Farré CA, Gerlach AL, Wittchen HU,

Lang T, Alpers GW, Fydrich T, Ströhle A, Arolt V, Schweiger S, Winter J, Mota NR, Franke B, Harneit A, Schweiger JI, Schwarz K, Ma R, Chen J, Schwarz E, Tost H, Meyer-Lindenberg A, Erk S, Heinz A, Romanczuk-Seiferth N, Walter H, Witt S, Rietschel M, Noethen MM, Richter J, Yang Y, Kircher T, Hamm AO, Straube B, Lueken U, Weber H, Deckert J, Freudenberg F, Cormand B, Slattery DA, Reif A

**Journal**: To be submitted

**Contribution of the PhD candidate**:

Analysis of GWAS data and enrichment analysis of *RBFOX1* targets. Comprehensive collection of data from published CNV studies in psychiatric disorders and search for CNVs described in *RBFOX1* in either patients or controls. Burden analysis for *RBFOX1* CNVs. Comprehensive collection of transcriptomic data from publications and GEO datasets. Preparing figures and tables and participating in manuscript edition.

**Article 4**:

**Title**: Pleiotropic contribution of *rbfox1* to psychiatric and neurodevelopmental phenotypes in a zebrafish model

**Authors**: Antón-Galindo E, Adel M, López-Blanch L, Norton WHJ, Fernàndez-Castillo N[#], Bru Cormand B[#]

[#] equally supervised this work

**Journal**: To be submitted

**Contribution of the PhD candidate**:

Crossing and genotyping zebrafish to obtain mutant lines. Collection of zebrafish brains and RNA extraction for both in situ hybridization (ISH) and qPCR. ISH of zebrafish larvae and adult brains. Quantification of *rbfox1* expression by qPCR. Setting up and performing the behavioural experiments and analyses with Python. Writing the first draft of the manuscript, figures and tables, and participating in the final edits.

**CHAPTER 3: Exploring the contribution of the *BEX/TCEAL* gene family to ASD and other psychiatric disorders**

<u>**Article 5**</u>:

**Title**: Characterization of an eutherian gene cluster generated after transposon domestication identifies *Bex3* as relevant for advanced neurological functions

**Authors**: Navas-Pérez E, Vicente-García C, Mirra S, Burguera D, Fernàndez-Castillo N, Ferrán JL, López-Mayorga M, Alaiz-Noya M, Suárez-Pereira I, <u>Antón-Galindo E</u>, Ulloa F, Herrera-Úbeda C, Cuscó P, Falcón-Moya R, Rodríguez-Moreno A, D'Aniello S, Cormand B, Marfany G, Soriano E, Carrión ÁM, Carvajal JJ, Garcia-Fernàndez J
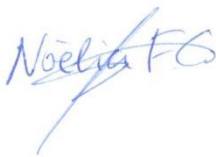
**Journal**: Genome Biology. 2020 Oct 26;21(1):267. doi: 10.1186/s13059-020-02172-3.

**Impact Factor (2019)**: 10.806
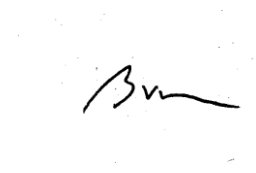
 **Contribution of the PhD candidate**:

Genotyping of almost 300 patients to investigate the presence of mutations in the *BEX3* gene. Comprehensive collection of transcriptomic data from publications and GEO datasets and subsequent enrichment analysis. Preparing tables and participating in manuscript edition.

Barcelona, 16 of June 2021

Noèlia Fernàndez Castillo                                          Bru Cormand Rifà

# CHAPTER 1.

## Exploring the contribution of the 14-3-3 gene family to ASD and other psychiatric disorders

# Article 1. Involvement of the 14-3-3 gene family in autism spectrum disorder and schizophrenia: genetics, transcriptomics and functional analyses

**Summary in Spanish: "Implicación de la familia de genes 14-3-3 en el trastorno del espectro autista y en esquizofrenia: genética, transcriptómica y análisis funcionales"**

La familia de proteínas 14-3-3 está compuesta por chaperonas moleculares implicadas en diferentes funciones biológicas y trastornos neurológicos. En un artículo previo, señalamos a *YWHAZ*, que codifica la proteína 14-3-3ζ, como gen candidato para el trastorno del espectro autista (TEA) mediante un estudio de secuenciación de exoma completo en el que se identificó una variante frameshift en este gen (c.659-660insT, p.L220Ffs*18). En este estudio hemos explorado la contribución de los siete miembros de la familia 14-3-3 al TEA y otros trastornos psiquiátricos analizando: i) el impacto funcional de la mutación p.L220Ffs*18 de la proteína 14-3-3ζ en su solubilidad, capacidad de unión a otra proteína y dimerización, ii) la contribución de variantes comunes de riesgo en los genes 14-3-3 al TEA y otros trastornos psiquiátricos, iii) la carga de variantes raras en los genes 14-3-3 en individuos con TEA y esquizofrenia, iv) la expresión alterada de los genes 14-3-3 utilizando datos transcriptómicos de pacientes con TEA y esquizofrenia. Demostramos que la proteína 14-3-3ζ mutada tiene una solubilidad inferior, es incapaz de formar heterodímeros y de unirse a la tirosina hidroxilasa. Además, mediante análisis genéticos usando datos públicos demostramos que variantes comunes en *YWHAE* contribuyen a esquizofrenia ($p$ = 6.6E-07) mientras que variantes ultra-raras en los genes 14-3-3 son más frecuentes en pacientes con TEA ($p$ = 0.016). Por último, la expresión de los genes 14-3-3 está alterada en cerebros *postmortem* de pacientes con TEA y esquizofrenia. Nuestro estudio sugiere por tanto un papel importante de la familia 14-3-3 en TEA y esquizofrenia.

**Reference:**

Torrico B, Antón-Galindo E, Fernàndez-Castillo N, Rojo-Francàs E, Ghorbani S, Pineda-Cirera L, Hervás A, Rueda I, Moreno E, Fullerton JM, Casadó V, Buitelaar JK, Rommelse N, Franke B, Reif A, Chiocchetti AG, Freitag C, Kleppe R, Haavik J, Toma C, Cormand B. Involvement of the 14-3-3 Gene Family in Autism Spectrum Disorder and Schizophrenia: Genetics, Transcriptomics and Functional Analyses. J Clin Med. 2020 Jun 13;9(6):1851.

*Article*

# Involvement of the 14-3-3 Gene Family in Autism Spectrum Disorder and Schizophrenia: Genetics, Transcriptomics and Functional Analyses

Bàrbara Torrico [1,2,3,4,†], Ester Antón-Galindo [1,2,3,4,†], Noèlia Fernàndez-Castillo [1,2,3,4,†], Eva Rojo-Francàs [1,2,3,4], Sadaf Ghorbani [5], Laura Pineda-Cirera [1,2,3,4], Amaia Hervás [6,7], Isabel Rueda [6], Estefanía Moreno [3,8], Janice M. Fullerton [9,10], Vicent Casadó [3,8], Jan K. Buitelaar [11,12], Nanda Rommelse [12,13], Barbara Franke [13,14], Andreas Reif [15], Andreas G. Chiocchetti [16], Christine Freitag [16], Rune Kleppe [5,17], Jan Haavik [5], Claudio Toma [1,9,10,18,*,‡] and Bru Cormand [1,2,3,4,*,‡]

1   Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Prevosti Building, floor 2, Av. Diagonal 643, 08028 Barcelona, Spain; barticoa@gmail.com (B.T.); eantongalindo@ub.edu (E.A.-G.); noefernandez@ub.edu (N.F.-C.); erojofra20@alumnes.ub.edu (E.R.-F.); l.pineda@ub.edu (L.P.-C.)
2   Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, 28029 Madrid, Spain
3   Institut de Biomedicina de la Universitat de Barcelona (IBUB), 08028 Barcelona, Spain; estefaniamoreno@ub.edu (E.M.); vcasado@ub.edu (V.C.)
4   Institut de Recerca Sant Joan de Déu (IR-SJD), 08950 Esplugues de Llobregat, Spain
5   Centre for Neuropsychiatric Disorders, Department of Biomedicine, University of Bergen, N5009 Bergen, Norway; Sadaf.Ghorbani@uib.no (S.G.); Rune.Kleppe@uib.no (R.K.); Jan.Haavik@uib.no (J.H.)
6   Child and Adolescent Mental Health Unit, Hospital Universitari Mútua de Terrassa, 08221 Terrassa, Spain; ahervas@mutuaterrassa.cat (A.H.); irueda@sjdhospitalbarcelona.org (I.R.)
7   IGAIN, Global Institute of Integral Attention to Neurodevelopment, 08007 Barcelona, Spain
8   Department of Biochemistry and Molecular Biomedicine, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain
9   Neuroscience Research Australia, Sydney, NSW 2031, Australia; j.fullerton@neura.edu.au
10   School of Medical Sciences, University of New South Wales, Sydney, NSW 2052, Australia
11   Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands; Jan.Buitelaar@radboudumc.nl
12   Karakter Child and Adolescent Psychiatry University Centre, 6525 GC Nijmegen, The Netherlands; Nanda.Lambregts-Rommelse@radboudumc.nl
13   Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands; Barbara.Franke@radboudumc.nl
14   Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, The Netherlands
15   Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt, 60590 Frankfurt am Main, Germany; andreas.reif@kgu.de
16   Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Autism Research and Intervention Center of Excellence Frankfurt, JW Goethe University, 60323 Frankfurt am Main, Germany; andreas.chiocchetti@kgu.de (A.G.C.); christinemargarete.freitag@kgu.de (C.F.)
17   Division of Psychiatry, Haukeland University Hospital, 5021 Bergen, Norway
18   Centro de Biología Molecular "Severo Ochoa", Universidad Autónoma de Madrid/CSIC, C/Nicolás Cabrera, 1, Campus UAM, 28049 Madrid, Spain
*   Correspondence: claudio.toma@cbm.csic.es (C.T.); bcormand@ub.edu (B.C.); Tel.: +34-911-964-731 (C.T.); +34-934-021-013 (B.C.)
†   These authors equally contributed to this work.
‡   These authors equally supervised this work.

check for updates

**Abstract:** The 14-3-3 protein family are molecular chaperones involved in several biological functions and neurological diseases. We previously pinpointed *YWHAZ* (encoding 14-3-3ζ) as a candidate gene for autism spectrum disorder (ASD) through a whole-exome sequencing study, which identified a frameshift variant within the gene (c.659-660insT, p.L220Ffs*18). Here, we explored the contribution of the seven human 14-3-3 family members in ASD and other psychiatric disorders by investigating the: (i) functional impact of the 14-3-3ζ mutation p.L220Ffs*18 by assessing solubility, target binding and dimerization; (ii) contribution of common risk variants in 14-3-3 genes to ASD and additional psychiatric disorders; (iii) burden of rare variants in ASD and schizophrenia; and iv) 14-3-3 gene expression using ASD and schizophrenia transcriptomic data. We found that the mutant 14-3-3ζ protein had decreased solubility and lost its ability to form heterodimers and bind to its target tyrosine hydroxylase. Gene-based analyses using publicly available datasets revealed that common variants in *YWHAE* contribute to schizophrenia ($p = 6.6 \times 10^{-7}$), whereas ultra-rare variants were found enriched in ASD across the 14-3-3 genes ($p = 0.017$) and in schizophrenia for *YWHAZ* (meta-$p = 0.017$). Furthermore, expression of 14-3-3 genes was altered in post-mortem brains of ASD and schizophrenia patients. Our study supports a role for the 14-3-3 family in ASD and schizophrenia.

**Keywords:** autism; 14-3-3 gene family; rare variants; common variants; transcriptomics; schizophrenia; *YWHAZ*; *YWHAE*

## 1. Introduction

Autism spectrum disorder (ASD) is characterized by impairments in communication and social interactions, and the presence of repetitive and restrictive behaviours [1]. However, the clinical picture is often accompanied by additional features, such as intellectual disability (ID), epilepsy, language impairment, anxiety, sleep disorders, and attention-deficit hyperactivity disorder (ADHD) [2]. ASD has a prevalence of approximately 1.5% in the general population [3,4] and large studies suggest an unequivocal genetic contribution to its aetiology. Indeed, family and twin studies indicate a heritability of around 80%, which represents one of the highest amongst neuropsychiatric disorders [5–7]. Despite the substantial role of genetic factors in the disorder, the genetic architecture is not fully dissected and many of the underlying genes are yet to be identified. Also, the genetic relationships amongst comorbid phenotypes remain largely unknown.

Genetic studies suggest a multi-hit model of inheritance in which a combination of risk alleles, including both common variants of small effect size and rare variants with higher penetrance, contribute to the phenotype [8,9]. Despite the large number of association studies in the last decade, the robust identification of common risk alleles for ASD has been elusive, and the first genome-wide association studies (GWAS) performed with reasonably large samples [10–12] pinpointed associations with single nucleotide polymorphisms (SNPs) that were subsequently not replicated in a large European sample [13]. A GWAS recently performed in 18,381 ASD patients and 27,969 controls identified five genome-wide significant loci [14], suggesting the need of large samples to identify common risk variants. Regarding rare variants in ASD, the first whole-exome sequencing (WES) studies focused on de novo variants and suggested a substantial role for this class of damaging mutations in the aetiology of the disorder [15–17]. These approaches were crucial to pinpoint novel genes involved in the disorder [18], and recently the largest WES study reported 102 ASD risk genes in a sample of 11,896 cases [19]. WES studies were also performed to address the impact of rare inherited variants [18,20–22]. We performed the first WES study in multiplex families with autism that suggested a role for genome-wide truncating mutations in the aetiology of ASD [22]. A current genetic model would implicate a higher number of gene-disrupting variants in severe ASD phenotypes, increasing symptom severity and comorbidity with ID [2,19].

In our first WES study [22] we identified a truncating mutation in the *YWHAZ* gene (c.659-660insT, p.L220Ffs*18) that was transmitted from a mother with depression to an ASD sib-pair. In that previous study, the *YWHAZ* gene was the main interconnected node in a network including mutated genes identified by previous WES studies and other genes implicated in ASD. This gene encodes 14-3-3ζ, a protein involved in a range of biological processes including cell cycle, transcription, neuronal development, migration and neurite outgrowth [23–25]. 14-3-3ζ is one of the seven members (β, γ, ε, η, ζ, σ, θ) of the highly conserved 14-3-3 protein family. These proteins exert their function as homo- and heterodimers through protein-protein interaction with a wide range of target proteins, typically binding to phosphorylated serine and threonine residues [26]. Several studies have demonstrated the role of *YWHAZ* and other family members in neurogenesis and neurodifferentiation, and its possible implication in neurodevelopmental disorders [24,27]. Recent studies suggested that decreased 14-3-3ζ protein levels in ASD may be responsible for deficits in melatonin synthesis observed in ASD via the downregulation of the aralkylamine N-acetyltransferase (AANAT) and acetylserotonin o-methyltransferase (ASMT) enzymatic activity [28–30]. Moreover, *YWHAZ* knock-out mice show impaired cortical development and larger lateral ventricles, reduced dendritic and synaptic density, aberrant neuronal migration in hippocampus, abnormal mossy fibers connectivity, and cognitive deficits [31–33].

Interestingly, several members of the 14-3-3 gene family have been implicated in psychiatric disorders, including reported associations of *YWHAB*, *YWHAE*, *YWHAZ* and *YWHAH* with schizophrenia [34–39]; *YWHAE* and *YWHAQ* with ADHD [39]; *YWHAH*, *YWHAG* and *YWHAE* with bipolar disorder [39,40]; *YWHAE* and *YWHAQ* with major depressive disorder [39,41]; and *YWHAE* with suicide attempts [42]. However, most of these studies were performed in small samples and lacked replication in larger populations. Furthermore, altered levels of 14-3-3 proteins were found in the blood or brains of patients with ASD [29,43,44], schizophrenia [45–48] and bipolar disorder [48]. Interestingly, microduplications of *YWHAE*, which encodes 14-3-3ε that form stable heterodimers with 14-3-3ζ, were reported in ASD patients; whereas microdeletions involving both *YWHAE* and *PAFAH1B1* genes cause Miller–Dieker syndrome, a form of lissencephaly with ID and seizures [49–53], and deletions including *YWHAG* and *HIP1* were related to epilepsy, learning problems and ID [54].

Animal models deficient for 14-3-3 proteins show a variety of behavioural manifestations related to psychiatric disorders: *Ywhaz* (14-3-3ζ) deficient mice present hyperactivity, impaired memory, lower anxiety and impaired sensorimotor gating [32,33]; *Ywhae* (14-3-3ε) deficient mice present enhanced anxiety-like behaviour, defects in working memory, increased locomotor activity and sociability [34,55]; and heterozygous knock-out mice deficient for *Ywhag* (14-3-3γ) show hyperactivity and depressive-like behaviour [56]. Furthermore, a recent study reported that the specific inhibition of 14-3-3 proteins in the hippocampus of mice is sufficient to cause hyperactivity, reduce sensorimotor gating and impair associative learning and memory [57].

Based on these aforementioned reports, we hypothesize that the 14-3-3 gene family may play an important role in the susceptibility to psychiatric disorders. Thus, here we aim to: (i) gain molecular insight for the role of the *YWHAZ* truncating mutation p.L220Ffs*18 previously reported in a family with distinct psychiatric disorders; (ii) assess the contribution of common and rare variants of the 14-3-3 gene family (*SFN*, *YWHAB*, *YWHAE*, *YWHAG*, *YWHAH*, *YWHAQ* and *YWHAZ*) to ASD and other psychiatric disorders; and (iii) explore possible altered expression levels of this gene family in psychiatric disorders.

## 2. Experimental Section

### 2.1. Expression, Purification and Solubility Testing of Recombinant 14-3-3ζ Wild-Type and Mutated Proteins

The recombinant human YWHAZ wild-type (WT) and mutant were expressed in *E. coli* using pGEX-2TK expression system, and purification of the soluble fractions as fusion proteins with glutathione-S-transferase (GST) was performed according to previous protocols [58], see also

supplementary information for details. The 14-3-3ζ mutated protein replicating the C-terminal amino acid alteration from the p.L220Ffs*18 insertion mutation, was generated using site-directed mutagenesis as detailed in the supplementary information.

Due to low solubility of the GST-14-3-3ζ_mut protein we tested expression and solubility at different temperatures (induction time, h) of 30 (4 h), 25 (5 h) and 20 (6 h) °C and compared total and soluble lysate to that of WT on sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE, see supplementary information for details).

## 2.2. Functional Assessment of 14-3-3ζ WT and Mutant Proteins by Surface Plasmon Resonance

One of the best-characterized binding targets of 14-3-3 proteins is tyrosine hydroxylase (TH) phosphorylated at serine 19 (THpSer19) [58,59]. Human TH was purified and phosphorylated on serine 19 using p38 regulated/activated protein kinase (PRAK, also referred to as MK2) as previously described [60]. The WT and mutated 14-3-3ζ proteins were assessed for their binding to THpS19 or non-phosphorylated TH using surface plasmon resonance (Biacore 3000, Cytiva, Marlborough, MA, USA) as previously described [58,60]. GST-14-3-3 proteins were immobilized on CM5 sensor chips according to manufacturer's instructions using the GST-capture kit (Cytiva) giving similar amounts of immobilized GST-14-3-3ζ_WT and GST-14-3-3ζ_mut. Target protein binding was assessed at 25 °C, using the HBS-P Buffer provided by the manufacturer (10 mM HEPES pH 7.4, 150 mM NaCl and 0.005% polysorbate 20). Different concentrations of Ser19 phosphorylated TH were injected multiple times for multiple immobilizations. We used the unphosphorylated TH as a negative control of the binding. The resulting sensograms were analyzed with the BIAevaluation v3.2 software (Biacore AB, Uppsala, Sweden).

## 2.3. Assessing the Dimerization of Mutant and Wild-Type Proteins (14-3-3ζ and 14-3-3σ) in Cells by Bioluminescence Resonance Energy Transfer (BRET) Assay

Plasmids were obtained for expressing fusion proteins of different 14-3-3 members (14-3-3ζ, 14-3-3ε, 14-3-3σ) with Rluc (Renilla luciferase, donor) and EYFP (enhanced yellow variant of GFP, acceptor), as described in the supplementary information. The ability of mutant 14-3-3ζ to form heterodimers with 14-3-3ε, and of mutant 14-3-3σ to form homodimers, was assessed using a bioluminescence resonance energy transfer (BRET) assay.

Human embryonic kidney cells (HEK-293T) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/mL penicillin and 100 μg/mL streptomycin (GIBCO, Carlsbad, CA, USA), in a 5% $CO_2$ humidified atmosphere at 37 °C. The cell line was grown in 6-well plates (35-mm-diameter wells) at a density of $5.0 \times 10^5$ cells/well for transfection using $CaCl_2$ as described in the supplementary information. Cells were co-transfected with the cDNA construct coding for Rluc-target-protein-1, acting as BRET donor, and increasing amounts of the cDNA construct coding for YFP-target-protein-2 as BRET acceptor (see Table S1 for the amounts of cDNA used). D(1A) dopamine receptor (DRD1) fusion protein with YFP or Rluc [61] was used as a negative control of the dimerization, and co-transfected with the corresponding tested constructs (Table S1). After 48 h upon transfection, cells were washed twice with Hanks' balanced salt solution pH 7.4 (HBSS, 137 mM NaCl, 5 mM KCl, 1.26 mM $CaCl_2$, 0.4 mM $MgSO_4$, 0.5 mM $MgCl_2$, 0.34 mM $Na_2HPO_4$, 0.44 mM $KH_2PO_4$, 10 mM HEPES) supplemented with 1% glucose (*w/v*), detached and resuspended in the same buffer. Protein concentration was determined using the Bradford assay kit (Bio-Rad, Munich, Germany) and all cell suspensions were diluted with HBSS to obtain a final concentration of 0.2 mg/mL of protein.

In order to quantify fluorescence, cell suspensions (20 μg of protein) were distributed in duplicates in a 96-well black microplate with a transparent bottom (Porvair, King's Lynn, UK). Fluorescence was then measured using a FLUOstar Optima fluorimeter (BMG Labtechnologies, Offenburg, Germany) equipped with a high-energy xenon flash lamp, and a 10 nm bandwidth excitation filter at 400 nm reading. A PHERAstar Flagship FSX fluorimeter (BMG Labtechnologies, Offenburg, Germany) was used for BRET and luminescence measurements. Cell suspensions (20 μg of protein) were distributed

in duplicates in a 96-well white opaque microplate (Porvair, King's Lynn, UK) and coelenterazine H (Molecular Probes Europe, Leiden, The Netherlands) was added at a final concentration of 5 mM. One minute after adding coelenterazine H, luminescence readings were collected using sequential integration of signals detected at 440–500 nm and 510–590 nm. Luminescence measurements of the same samples were performed after 10 min of incubation with coelenterazine H. Cells expressing BRET donors alone were used to determine background. The BRET ratio is defined as: ((emission at 510–590 nm)/(emission at 440–500 nm))–Cf; where Cf corresponds to (emission at 510–590 nm)/(emission at 440–500 nm) for the donor construct expressed alone in the same experiment. Data were fitted to a non-linear regression equation, assuming a single phase saturation curve with GraphPad Prism software (San Diego, CA, USA).

*2.4. Common Variants in the 14-3-3 Family in Autism Spectrum Disorder (ASD): Association Study in Our Sample*

2.4.1. Subjects of Our ASD Cohorts

The cohort used in the case-control association study consisted of 727 ASD patients and 714 gender-matched controls from three European populations: Spanish, Dutch and German (Table S2). The cohort used for mutational screening employed a subset of 288 ASD patients from the same three populations (Table S2). All individuals had European ancestry. ASD patients were assessed using the ADI-R (Autism-Diagnostic Interview-Revised) [62] and, where possible, also the ADOS (Autism Diagnostic Observation Schedule) [63]. Cytogenetic abnormalities or a positive Fragile X test were considered exclusion criteria. The study was approved by the relevant ethics committee from each center and written informed consent was obtained from all parents/guardians or, where possible, by affected individuals, according to the Helsinki Declaration. Genomic DNA was extracted from peripheral blood samples using the standard salting-out method [64].

2.4.2. Common Variant Association Study of the 14-3-3 Family in Our ASD Sample

SNP selection was performed to encompass common genetic variants across the seven genes of the 14-3-3 family (*SFN, YWHAQ, YWHAG, YWHAZ, YWHAB, YWHAH* and *YWHAE*). Each gene included a 5kb flanking region (both 5′ and 3′), and patterns of linkage disequilibrium (LD) were considered using the Central European (CEU) panel of HapMap project data (www.hapmap.org; phases 1, 2, 3; release 28). A total of 42 tagSNPs were selected using the Tagger implementation in HaploView v4.2 [65], according to the following criteria: $r^2 \leq 0.8$ and minor allele frequency (MAF) $\geq 0.05$. The sample of 1441 subjects was genotyped using iPlex-Sequenom technology (Sequenom, San Diego, CA, USA) at the Spanish National Genotyping Center (CeGen). Duplicates were included as controls for genotyping quality. After quality control procedures, 36 individuals with genotyping rate lower than 90% were removed from the study, setting the final sample to 1405 individuals (713 cases and 692 controls). From the 42 SNPs initially genotyped, four assays failed, and one SNP was monomorphic. The genotyping rate for the 37 remaining SNPs was 99.2%. One of the SNPs was excluded for quality reasons, but all the remaining 36 SNPs were in Hardy–Weinberg equilibrium (threshold set at $p < 0.01$ in controls). Thirty-four SNPs showed a MAF $> 0.05$ in our sample and were examined for association with ASD (Table S3). LD patterns and D′ values were determined in our sample data with Haploview v4.2 [65] (Figure S1). A case-control association study under the additive model was run with the PLINK package [66].

*2.5. Common Variant Associations with the 14-3-3 Gene Family across Psychiatric Disorders Using Public Genome-Wide Association Studies (GWAS) Data*

We assessed the contribution of common variation in the seven 14-3-3 genes to psychiatric disorders using GWAS summary statistics from the Psychiatric Genomics Consortium (PGC), Broad Antisocial Behaviour Consortium (BroadABC) and Integrative Psychiatric Research Consortium (iPSYCH). We considered the following phenotypes: attention-deficit/hyperactivity disorder (ADHD) [67], anti-social behaviour (ASB) [68], anxiety [69], autism spectrum disorder (ASD) [14], bipolar disorder (BD) [70], major depressive disorder (MDD) [71], obsessive-compulsive disorder (OCD) [72], schizophrenia [73] and cross-disorder meta-analysis [74] (see details in Table S4). All summary statistics used for subsequent analysis had a MAF ≥ 0.01 and info-score for imputation quality ≥ 0.6. A gene-based association study was performed with MAGMA (v1.06) [75] using the 1000 Genomes Project Phase 3 (Build 37/European data only) as a reference panel. We also performed a self-contained gene-set analysis considering the whole 14-3-3 gene family for each of the eight psychiatric phenotypes and for the cross-disorder meta-analysis.

*2.6. Rare Variant Analysis of the 14-3-3 Gene Family: Mutational Screening in Our ASD Sample*

The seven genes of the 14-3-3 family were analyzed in 288 Caucasian ASD patients by high-throughput sequencing using the Ion Torrent platform (ThermoFisher Scientific, Waltham, MA, USA) at the Centre for Research in Agricultural Genomics (CRAG). A total of 57 tagged-primer pairs were designed with the Ion Ampliseq Designer (ThermoFisher Scientific) and covered 96.3% of all coding exons across the 14-3-3 gene family (Table S5), including the splice sites and part of the 5′ and 3′ untranslated regions (UTR) (Figure S2). The corresponding amplicons were sequenced in 288 ASD patients (182 Spanish, 94 Dutch, 12 Germans), including the MT_160.3 patient, heterozygous carrier for the previously described c.659_660insT mutation leading to p.L220Ffs*18 in the *YWHAZ* gene [22], as a positive control. Genomic DNA was pooled in groups of three subjects to minimize costs. The results were processed using the Ion Reporter Software (ThermoFisher Scientific) under somatic variant calling parameters to identify low-frequency variant calls (average read depth 1400X), and the function Variant Analysis of the Ingenuity Pathway Analysis software (http://www.ingenuity.com/products/ipa) was employed to assess the identification and the molecular nature of the variants identified in our sample. To handle analysis of pooled sequencing, every single change identified in a pool was subjected to Sanger validation in the three subjects present in the reaction, in order to confirm the variant and determine the carrier status of each individual.

*2.7. Rare Variant Analysis of the 14-3-3 Gene Family: ASD and Schizophrenia Public Datasets*

The impact of rare variants across the seven genes of the 14-3-3 family was assessed and further extended using publicly available sequencing data of schizophrenia, ASD and control cohorts from the following sources: (i) whole-exome sequencing (WES) from the Sweden-Schizophrenia population-based case-control (database of Genotypes and Phenotypes (dbGAP) accession: phs000473.v2.p2) (6135 cases and 6245 controls); (ii) ARRA Autism Sequencing Collaboration (dbGAP accession: phs000298.v3.p2) (1288 unrelated ASD probands); (iii) European ASD samples, which included our previously sequenced sample of 288 ASD patients plus 348 additional ASD patients from Germany; (iv) Medical Genome Reference Bank (2845 healthy Caucasian Australians aged > 75). The selection of variants was based on: (1) their predicted pathogenicity using the Variant Effect Predictor annotation tool software (https://www.ensembl.org/Tools/VEP): missense mutations predicted to be damaging in both SIFT (Sorting Intolerant From Tolerant) and Polyphen2 (Polymorphism Phenotyping v2)), and with CADD (Combined Annotation Dependent Depletion) > 20 for canonical splice site variants, stop codon mutations and indels leading to frameshift; and (2) minor allele frequency (MAF < 0.0001) in non-Finnish European populations from the Genome Aggregation Database (http://gnomad.broadinstitute.org/) as previously described [18]. A chi-square statistic was used to compare the schizophrenia patient sample (6135 cases) and combined ASD datasets (1924 cases)

with the combined control datasets (9090 individuals). Additionally, we explored the impact of rare variants for each of the 14-3-3 genes as reported by the Autism Sequencing Consortium (ASC) (https://asc.broadinstitute.org/) and the Schizophrenia Exome-sequencing Meta-Analysis (SCHEMA) (https://schema.broadinstitute.org/). ASC represents the largest source of data of rare variants for genetic studies of ASD, which includes de novo variant calls for 6430 probands and 2179 unaffected siblings (family-based dataset), and rare variants identified in 5556 ASD patients and 8809 controls (case-control dataset). Similarly, SCHEMA combines data for rare variants across several world-wide populations with a joint sample of 24,248 schizophrenia patients and 97,322 controls (case-control dataset), and 3444 schizophrenia trios (family-based dataset).

### 2.8. Expression of the 14-3-3 Genes in Autism Spectrum Disorder and Schizophrenia

Differential expression of the seven 14-3-3 genes was assessed using transcriptomic data from post-mortem brain regions of ASD and schizophrenia patients in publicly available human datasets, either in the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) or in published articles. We found 39 studies with available information on gene expression in brain: 12 in ASD patients and 27 in schizophrenia patients. In particular, in the case of ASD we analysed data on altered gene expression in brain from 11 papers and one GEO dataset (PubMed ID: 27919067, 29859039, 21614001, 22457638, 22984548, 18006270, 18378158, 25494366, 27219343, 27685936, 30545856 and GSE38322). For schizophrenia we analysed data on brain gene expression from 14 papers (30545856, 25113377, 24287731, 21091092, 24167345, 23904455, 24686180, 24886351, 22031440, 18778695, 26818902, 22212594, 22954356, 29931221) and 13 GEO datasets (GSE46509, GSE37981, GSE21935, GSE21138, GSE17612, GSE12654, GSE87610, GSE53987, GSE62191, GSE35977, GSE12649, GSE12679, GSE35978). Differential expression was assessed in multiple brain areas, including hippocampus, cerebellum or cortex, depending on the dataset.

### 3. Results

#### 3.1. Functional Effect of the YWHAZ (14-3-3ζ) Mutation p.L220Ffs*18

We firstly aimed to assess the functional consequences of a truncating variant (c.659_660insT, p.L220Ffs*18) previously identified in the *YWHAZ* gene and transmitted from a mother with depression, phobia and fibromyalgia to two siblings both with ASD and attention-deficit hyperactivity disorder (ADHD) [22]. The maternal grandmother was diagnosed with schizophrenia (Figure 1A), although DNA from this case was unfortunately not available for mutation analysis.

This truncating variant does not meet criteria to trigger degradation of the corresponding mRNA by nonsense-mediated RNA decay (NMD), since it is located 22 nucleotides before the last exon-exon junction (Figure 1B) [76,77]. Therefore, we performed experiments to investigate the possible functional effect of p.L220Ffs*18 on the protein by assessing: (i) its solubility; (ii) its ability to bind tyrosine hydroxylase (TH), one of its natural targets when phosphorylated at its 14-3-3 binding site (THpSer19); and (iii) its capacity to form heterodimers with 14-3-3ε.

The mutated 14-3-3ζ protein (p.L220Ffs*18) showed decreased solubility compared to the WT 14-3-3ζ when it was expressed in prokaryotes (*E. coli*) (Figure 1C,D). The solubility did not show any improvement at different temperatures for the mutated protein (Figure 1C). The fraction of soluble mutated protein was lower than 5% (compared to 90% for the WT protein), with the vast majority remaining insoluble in the lysate fraction (Figure 1D).
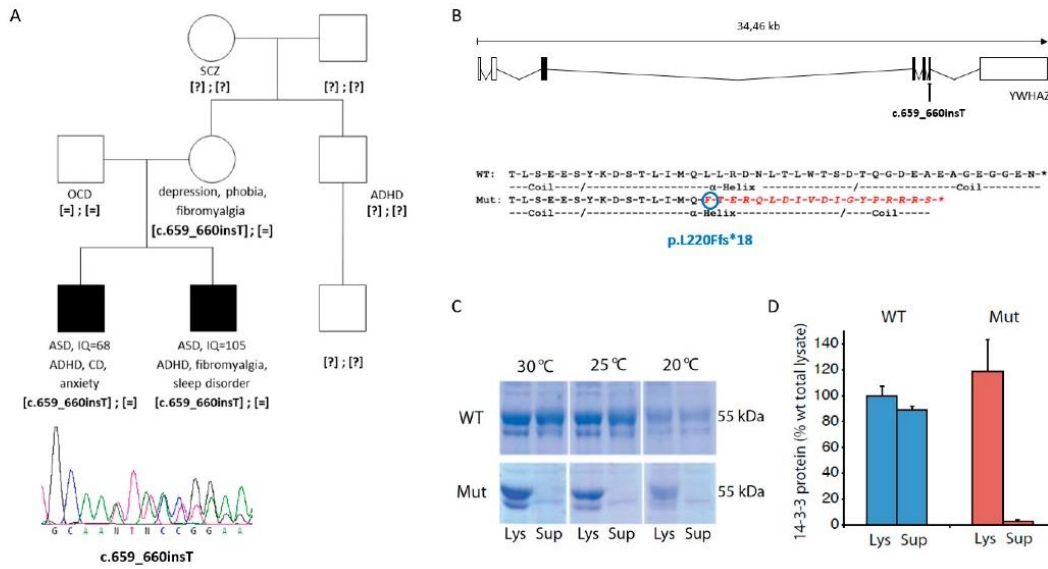
**Figure 1.** Mutation identified in *YWHAZ* in two siblings with ASD. (**A**) Pedigree of the family carrying the c.659-660insT mutation in *YWHAZ* and below the Sanger sequence of the truncating variant. Abbreviations: ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; CD, conduct disorder; IQ, intelligence quotient; OCD, obsessive-compulsive disorder; SCZ, schizophrenia. [=], wild-type allele; [?], unknown genotype. (**B**) Location of the mutation in the *YWHAZ* gene, and comparison of the last amino acids of the wild-type (WT) and mutant (Mut) 14-3-3ζ protein, showing in red the amino acids that diverge in the mutant protein. (**C**) Coomassie stained sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) of glutathione-S-transferase (GST)-14-3-3ζ (WT and mutant) expressed in *E. coli* strain at different temperatures. Aliquots of the bacteria were lysed, and the amount of soluble protein was assessed by comparing total lysate (Lys) to supernatant after centrifugation of the lysate (Sup) for wild-type 14-3-3ζ (WT) and mutant 14-3-3ζ (Mut). (**D**) Quantification of solubility by measuring the major GST-14-3-3 band (55 kDa) for WT and Mut expressed at 30 °C. Data presented as means and error bars denote the standard deviation ($n = 3$; $p = 1.7 \times 10^{-3}$ for Mut Lys vs. Mut Sup; $p = 1.4 \times 10^{-4}$ for WT Sup vs. Mut Sup; t-test, two sided).

The 14-3-3 proteins exert their function as homo- or heterodimers through binding to their target proteins, usually in a Ser/Thr phosphorylation-dependent manner [78]. We assessed the ability of 14-3-3ζ p.L220Ffs*18 to bind tyrosine hydroxylase (TH), the rate limiting enzyme in the synthesis of dopamine and one of its canonical target proteins. WT 14-3-3ζ has previously been reported to bind with nM affinity to Ser19-phosphorylated TH (THpSer19) (THpSer19) [58,79]. By using surface plasmon resonance, we found that the mutant protein expressed in prokaryotes (*E. coli*) lost its ability to bind the THpSer19 compared to the WT protein (Figure 2A,B).

We further assessed the capacity of both 14-3-3ζ_WT and 14-3-3ζ_mut proteins to interact with 14-3-3ε and form heterodimers, all expressed in a human cell line (HEK 293T cells). Through BRET assays we were able to detect the interaction between WT 14-3-3ζ and 14-3-3ε, as shown by a positive and saturable BRET signal (Figure 2C). Our results showed that the mutant 14-3-3ζ lost its capacity to form heterodimers with 14-3-3ε, since the linear relationship with the acceptor/donor ratio (YFP/Rluc) suggested lack of interaction (Figure 2C). We used DRD1, not known to be a 14-3-3 target, fused either to luciferase or to yellow fluorescent protein (DRD1-Rluc and DRD1-YFP) as negative controls. No interactions were identified, obtaining linear non-specific BRET signals, confirming the specificity of the WT 14-3-3ζ and 14-3-3ε interaction (Figure S3). Remarkably, when expressed in HEK293T cells for BRET assays, the 14-3-3ζ mutant protein only showed a detectable signal when transfecting at least 10 times the amount of plasmid required for the WT (Table S1), indicating likely protein degradation, in line with the decreased solubility observed in *E. coli*.
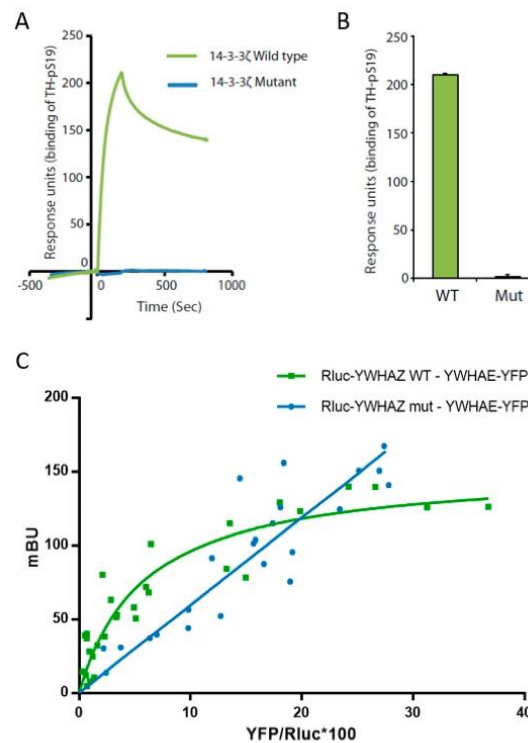
**Figure 2.** Characterization of mutant 14-3-3ζ binding capacity. (**A**) Binding of Ser19 phosphorylated human tyrosine hydroxylase (25 nM) to immobilized wild-type GST-14-3-3ζ (green) or mutant GST-14-3-3ζ (blue) using surface plasmon resonance (Biacore 3000, see Materials and Methods for details). The proteins were expressed in the BL21 Codon Plus *E. coli* strain and purified prior to the analysis. (**B**) The binding response for Ser19 phosphorylated tyrosine hydroxylase (25 nM) at the end of the injection was compared between 14-3-3ζ WT or Mut and Ser19 phosphorylated tyrosine hydroxylase (TH-pS19) (25 nM) repeated by several immobilizations and injections. Data presented as means and error bars denote the standard deviation ($n = 3$, $p = 1.61 \times 10^{-8}$ WT vs. Mut, two-sided t-test). (**C**) Characterization of wild-type 14-3-3ζ (YWHAZ WT) and mutant 14-3-3ζ (YWHAZ mut) interaction with 14-3-3ε (YWHAE) using a bioluminescence resonance energy transfer (BRET) assay. Rluc-YWHAZ WT co-transfection with an increasing amount of YWHAE-YFP gives a saturable positive signal (green), whereas the signal obtained co-transfecting Rluc-YWHAZ mut with an increasing amount of YWHAE-YFP fits with a linear regression (blue). mBU, BRET ratio expressed in milli-BRET units. The relative amount of BRET is given as a function of YFP/Rluc*100, where YFP corresponds to the fluorescence signal due to the increasing amount of donor and Rluc corresponds to the stable luminescence signal measured at 10 min. Values shown correspond to independent experiments ($n = 4$).

Thus, a damaging effect of the truncating mutation p.L220Ffs*18 on 14-3-3ζ was confirmed, with loss of function of the altered protein.

*3.2. Common Variants Across the 14-3-3 Gene Family in ASD and Other Psychiatric Disorders*

We sought for common genetic risk variants in the gene family encoding the 14-3-3 proteins in ASD and in other psychiatric disorders.

A case-control study was first performed in our sample of 713 ASD patients and 692 controls, both with European ancestry, investigating the common genetic variability of the 14-3-3 gene family (*SFN, YWHAB, YWHAE, YWHAG, YWHAH, YWHAQ* and *YWHAZ*) tagged by 34 SNPs. Only the variant rs1883660, located at the 3′UTR of the *SFN* gene, showed a nominal association with ASD ($p = 0.01$), but it did not remain significant after Bonferroni correction for multiple testing (Table S3).

We further extended the analysis for contribution of common variants in the seven genes of the 14-3-3 family to large GWAS datasets of eight psychiatric disorders, using summary statistics from the PGC, BroadABC and iPSYCH GWAS datasets (Table S4). The gene-based association study showed nominal associations for four out of seven genes with several psychiatric disorders: *YWHAB* with ADHD and schizophrenia, *YWHAE* with bipolar disorder and schizophrenia, *YWHAZ* with MDD and schizophrenia, and *SFN* with anxiety (Table 1).

**Table 1.** Gene-based association analysis of each of the seven 14-3-3 genes across several psychiatric disorders.

| Gene Symbol | Entrez ID | ADHD | ASB | Anxiety | ASD | BD | MDD | OCD | SCZ | Cross-Disorder |
|---|---|---|---|---|---|---|---|---|---|---|
| *YWHAB* | 7529 | **0.024** | 0.923 | 0.098 | 0.711 | 0.629 | 0.777 | 0.170 | **0.001** | 0.050 |
| *YWHAE* | 7531 | 0.086 | 0.933 | 0.190 | 0.468 | **0.006** | 0.154 | 0.262 | **$1.35 \times 10^{-6}$** | **$1.01 \times 10^{-5}$** |
| *YWHAG* | 7532 | 0.081 | 0.054 | 0.111 | 0.929 | 0.065 | 0.854 | 0.454 | 0.119 | 0.169 |
| *YWHAH* | 7533 | 0.153 | 0.520 | 0.218 | 0.741 | 0.104 | 0.658 | 0.243 | 0.291 | 0.410 |
| *YWHAQ* | 10971 | 0.801 | 0.438 | 0.927 | 0.835 | 0.637 | 0.441 | 0.054 | 0.300 | 0.900 |
| *YWHAZ* | 7534 | 0.262 | 0.415 | 0.311 | 0.125 | 0.168 | **0.029** | 0.771 | **0.001** | 0.095 |
| *SFN* | 2810 | 0.834 | 0.767 | **0.045** | 0.272 | 0.648 | 0.849 | 0.375 | 0.979 | 0.966 |

*p*-values were calculated using MAGMA (v1.06) software. Nominal associations are highlighted in bold. Underlined values survived Bonferroni correction for multiple testing, $p = 7.9 \times 10^{-4}$ (7 genes and 9 phenotypes). ADHD: attention-deficit hyperactivity disorder; ASB: antisocial behaviour; ASD: autism spectrum disorder; BD: bipolar disorder; MDD: major depression disorder; OCD: obsessive-compulsive disorder; SCZ: schizophrenia.

However, the association found between *YWHAE* and schizophrenia ($p = 1.35 \times 10^{-6}$; 33,640 cases and 43,456 controls) was the only surviving Bonferroni correction for 7 genes and 9 phenotypes. The gene-based results from the cross-disorder meta-analysis combining data across eight psychiatric disorders also showed a significant association with *YWHAE* ($p = 1.01 \times 10^{-5}$). We also performed a combined gene-set association analysis that indicates a nominal association of the whole 14-3-3 gene family with schizophrenia ($p = 0.018$) (Table S6).

### 3.3. Rare Variants in the 14-3-3 Gene Family in ASD and Schizophrenia

We investigated the role of rare variants in the 14-3-3 gene family in ASD and schizophrenia. First, we used next-generation sequencing to search for rare variants in the seven 14-3-3 genes in 288 ASD patients from our European collection, identifying nine rare variants (Table S7). All variants were validated by Sanger sequencing and the parental origin was assessed when possible (Table S7). Two detected variants were predicted to be deleterious and were found in the *SFN* gene in the same patient (p.E75del and p.T165S) together with a third variant predicted as benign (p.S149L). All three variants were present on the same chromosome and transmitted from the mother to an affected ASD proband (Figure S4A,B). We subsequently assessed the functional effect of 14-3-3σ carrying these three changes using BRET assays, observing that the mutant protein was still able to interact with both the WT and the mutated 14-3-3σ protein, as shown by positive saturable signals with similar $BRET_{50}$ and BRETmax (Figure S4C). The specificity of this interaction was confirmed by negative controls, obtaining linear non-specific BRET signals (Figure S4D). These results suggest that these rare variants do not impact dimerization.

We also explored the impact of ultra-rare variants (URV) across the 14-3-3 gene family in an extended sample of ASD patients, as well as in a sample of schizophrenia patients. For this objective we used publicly available sequencing data, which comprised 1924 ASD probands, 6135 schizophrenia patients, and 9090 control individuals. The high degree of evolutionary conservation of the seven 14-3-3 genes [80,81] and the limited sample size, resulted in relatively limited numbers of genetic variants. This prompted us to combine the data for the whole gene family. Interestingly, a significant burden of URVs was observed for ASD in the 14-3-3 family when compared to controls (7 in 1924 ASD cases vs. 11 in 9090 controls, $p = 0.017$), driven by a splice site variant in *YWHAE* (rs756213490), which was found in four unrelated ASD probands and not in controls (Table 2).

**Table 2.** List of ultra-rare variants (URVs) across the 14-3-3 gene family found in large public datasets.

| Chr:position | Ref/Alt | Gene | Data_Set (Controls/Cases) | gnomAD NFE AF | Impact | Amino Acids | SIFT | PolyPhen-2 | CADD | Existing Variation |
|---|---|---|---|---|---|---|---|---|---|---|
| 17:1264594 | T/A | YWHAE | ARRA (0/4) | $2.19 \times 10^{-5}$ | splice_acceptor_variant | canonical splice site | - | - | 34 | rs756213490 |
| 8:101961051 | C/A | YWHAZ | German_ASD (0/1) | | missense variant | A/S | deleterious low_confidence (0.03) | possibly_damaging (0.799) | 26.9 | rs774415799 |
| 1:27189925 | GGA/- | SFN | Spanish ASD (0/1) | $2.65 \times 10^{-4}$ | inframe deletion | SL/S | - | - | 21.5 | rs773116730 |
| 8:101936203 | A/A1 | YWHAZ | Spanish ASD (0/1) | | frameshift_variant | frameshift | - | - | - | - |
| 2:9731521 | G/A | YWHAQ | Swedish_SCZ (0/1) | $8.81 \times 10^{-6}$ | stop gained | Q* | - | - | 42 | rs769768541 |
| 8:101936511 | G/C | YWHAZ | Swedish_SCZ (0/1) | $8.84 \times 10^{-6}$ | stop gained | S* | - | - | 40 | rs754522887 |
| 22:32352395 | CAAGGTGTTTTACCTCGA/C | YWHAH | Swedish_SCZ (0/1) | $8.80 \times 10^{-6}$ | frameshift variant | KVLYLK/X | - | - | 35 | rs759167778 |
| 1:271189840 | T/C | SFN | MRGB (1/0) | $2.64 \times 10^{-5}$ | missense_variant | V/A | deleterious low_confidence (0.01) | probably_damaging (0.954) | 28.6 | rs77608477 |
| 1:27190047 | A/C | SFN | MRGB (1/0) | | missense_variant | I/A | deleterious low_confidence (0) | probably_damaging (0.929) | 26.5 | - |
| 1:27190145 | C/T | SFN | MRGB (1/0) | | missense_variant | R/W | deleterious low_confidence (0.03) | possibly_damaging (0.776) | 33 | - |
| 2:9725474 | G/GTGTTTAGG TTGT | YWHAQ | MRGB (1/0) | | frameshift_variant | frameshift | - | - | - | - |
| 8:101961101 | A/C | YWHAZ | MRGB (1/0) | | missense_variant | I/R | deleterious low_confidence (0.01) | probably_damaging (0.909) | 27.4 | - |
| 22:32352237 | T/C | YWHAH | MRGB (1/0) | $8.79 \times 10^{-6}$ | missense variant | V/A | deleterious low_confidence (0) | probably_damaging (0.999) | 27.9 | rs1196036662 |
| 22:32352724 | A/G | YWHAH | MRGB (2/0) | | missense_variant | N/S | deleterious low_confidence (0) | probably_damaging (0.991) | 26.5 | - |
| 1:27189780 | T/G | SFN | MRGB (1/0) | $8.80 \times 10^{-6}$ | missense variant | M/R | deleterious low_confidence (0.01) | possibly_damaging (0.631) | 27.7 | rs747687239 |
| 1:27190358 | CTG/C | SFN | Swedish_control (1/0) | $8.81 \times 10^{-6}$ | frameshift variant | frameshift | - | - | 35 | rs774524068 |
| 17:1264594 | T/A | YWHAE | Swedish_control (1/0) | $2.19 \times 10^{-5}$ | splice acceptor variant | canonical splice site | - | - | 34 | rs756213490 |

ARRA_c1, ASD cases from ARRA c1 data set (dbGAP accession: phs000298.v3.p2); German_ASD, ASD samples from Germany; MGRB, Medical Genome Reference Bank; Spanish_ASD, ASD samples from Spain; Swedish_SCZ, Sweden-Schizophrenia population-based Case-Control (dbGAP accession: phs000473.v2.p2);. All the URVs variants are selected to be rare (MAF < 0.0001 in Non-Finish European population in gnomAD, https://gnomad.broadinstitute.org/) and predicted to be pathogenic both in SIFT and, and CADD > 20.

No significant burden was identified for schizophrenia (3 in 6135 cases vs. 11 in 9090 controls, $p = 0.71$). It is noteworthy that truncating variants in *YWHAZ* are reported here only in patients (i.e., the frameshift c.659_660insT found in the ASD family functionally investigated here, and a stop mutation rs754522887 found in a Swedish schizophrenia patient), and they were not observed in the control group. Furthermore, we also explored large available datasets of ASD and schizophrenia from the Broad Institute for enrichment of rare variants, and observed that when the 14-3-3 genes were considered individually, only *YWHAZ* reached significance for a higher number of SNVs in schizophrenia (meta-analysis $p = 0.017$).

### 3.4. Altered Expression of the 14-3-3 Genes in ASD and Schizophrenia

Finally, we explored potential alterations of expression of the 14-3-3 genes in post-mortem brains of ASD and schizophrenia patients using transcriptomic datasets or literature reports. From the 39 studies with available information on gene expression in the brain (12 in ASD patients and 27 in schizophrenia patients), we found 11 studies reporting significant altered expression in 14-3-3 genes ($p < 0.05$) in ASD or schizophrenia. Seven of them reported altered 14-3-3 genes expression with a false discovery rate (FDR) < 0.1.

We found altered expression of six of the seven 14-3-3 genes in at least one of the two phenotypes compared to control subjects ($p < 0.05$, FDR < 0.1) (Table 3). All six genes, except for *SFN*, showed a decreased expression in different brain areas. Interestingly, five of these genes showed alterations in expression in both ASD and schizophrenia: *YWHAB*, *YWHAE*, *YWHAH* and *YWHAZ* showed decreased expression whether *SFN* showed increased expression in both disorders. In the case of *YWHAQ*, we found decreased expression only in ASD patients compared to controls (Table 3).

**Table 3.** Altered expression of the 14-3-3 genes in individuals with schizophrenia or autism spectrum disorder.

| Gene Symbol | Disorder | FC [&] | *p*-Value | FDR | Probe | Tissue | Study (PMID) or GEO ID | Sample, Cases vs. Controls |
|---|---|---|---|---|---|---|---|---|
| *YWHAB* | SCZ | −1.06 | $2.51 \times 10^{-3}$ | 0.10 | 8062880 | cerebellum | GSE35978 | 44 SCZ vs. 50 control |
| *YWHAB* | SCZ | −1.41 | 0.001 | 0.02 | 217717_s_at * | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *YWHAB* | ASD | N/A | N/A | $1.43 \times 10^{-3}$ | N/A | cortex (BA19, BA10, BA44) | Gupta et al., 2014 (25494366) | 32 ASD vs. 40 control |
| *YWHAE* | ASD | −1.34 | 0.003 | 0.06 | ILMN_1807535 | cerebellum | GSE38322 | 15 ASD vs. 12 control |
| *YWHAE* | SCZ | −1.52 | $1.93 \times 10^{-4}$ | 0.01 | 210317_s_at | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *YWHAE* | SCZ | −1.10 | $3.31 \times 10^{-4}$ | 0.04 | 11753092_s_at * | DLPFC | GSE87610 | 65 SCZ vs. 72 control |
| *YWHAH* | SCZ | −1.68 | $6.69 \times 10^{-5}$ | 0.01 | 201020_at | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *YWHAH* | ASD | N/A | <0.01 | <0.078 | N/A | DLPFC | Liu et al., 2016 (27685936) | 34 ASD vs. 40 control |
| *YWHAQ* | SCZ | −1.02 | 0.004 | 0.03 | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 560 SCZ vs. 936 control |
| *YWHAQ* | SCZ | −1.34 | $2.34 \times 10^{-5}$ | 0.01 | 200693_at * | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *YWHAZ* | ASD | −1.61 | $3.01 \times 10^{-5}$ | 0.01 | ILMN_1669286 | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| *YWHAZ* | SCZ | −1.87 | $3.07 \times 10^{-5}$ | 0.01 | 200641_s_at * | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *SFN* | ASD | 2.18 | 0.001 | 0.03 | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 51 ASD vs. 936 control |
| *SFN* | SCZ | 1.34 | $7.69 \times 10^{-5}$ | $1.33 \times 10^{-3}$ | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 559 SCZ vs. 936 control |
| *SFN* | SCZ | 1.53 | 0.010 | 0.08 | 33323_r_at * | hippocampus | GSE53987 | 15 SCZ vs. 18 control |

ASD, autism spectrum disorder; SCZ, schizophrenia; FDR, false discovery rate; [&] FC, fold change was calculated when log2FC was provided in the study; N/A, data not available; * Genes showing significant differential expression in independent probe sets targeting different transcripts/exons of the same gene, data shown corresponding to the probe with the highest fold-change.

## 4. Discussion

The 14-3-3 gene family encodes seven proteins that act as effectors of signaling-regulated proteins. They are highly expressed in the brain during development [82], and are involved in several neuronal processes, such as differentiation, migration, synaptogenesis, and axon guidance [24,27,31], as well as metabolic regulation [83]. Their important role in neuronal functions makes them plausible candidate genes for ASD and other psychiatric disorders. In a previous WES study we identified a truncating mutation in *YWHAZ* (encoding 14-3-3ζ), one of the members of the 14-3-3 family, present in two brothers with ASD [22]. Both individuals were diagnosed also with ADHD. One proband had mild intellectual disability (IQ of 68) and presented severe conduct disorder, aggressive behaviour and anxiety. The other proband had a normal IQ (105) and presented fibromyalgia and sleep disorder. Furthermore, in this family the mother of the ASD sib-pair, carrier of the *YWHAZ* mutation, presented with depression and other conditions, including phobia, fibromyalgia, hypothyroidism, asthma and obesity. The maternal grandmother of the ASD sib-pair was diagnosed with schizophrenia, and a maternal uncle of the two sibs had ADHD. Unfortunately, the potential segregation of this truncating variant in *YWHAZ* in the broader family could not be tested, as DNA collection was not possible. Interestingly, genes involved in dopamine neurotransmission have been reported to be associated with ADHD and ASD [84–86], and fibromyalgia and chronic pain have been related to decreased dopaminergic activity [87,88]. The 14-3-3 proteins act as regulators of dopamine synthesis by binding and stabilizing tyrosine hydroxylase [89]. Our study showed that the mutated protein 14-3-3ζ (p.L220Ffs*18) was not able to bind the phosphorylated tyrosine hydroxylase, a well-established molecular partner, leading likely to altered dopamine synthesis. Moreover, the mutated protein presented a decreased solubility when expressed in *E. coli* and it was not able to form heterodimers with 14-3-3ε when expressed in a human cell line. Our results are interesting also from the perspective of understanding the structural basis of 14-3-3 protein functions, as they suggest that the far C-terminal region may be involved in dimerization and phospho-target interaction, although further experiments would be needed to confirm it.

Previous research indicates that 14-3-3ζ deficient mice (*Ywhaz* knock-out) display cognitive and behavioural deficits possibly related to the dopamine system, altered hippocampal development, defective migration of pyramidal and granular neurons [33,90]. Also, the *Ywhaz* and *Ywhae* double knock-out mice show impaired neurogenesis, neuronal proliferation, differentiation and migration, and present severe seizures, indicating that both 14-3-3ζ and 14-3-3ε play a critical role during cortical development [31]. Indeed, 14-3-3 inhibition in certain brain regions in mice leads to impaired learning, working memory and long-term synaptic plasticity, symptoms that are associated with schizophrenia-like behaviours [91]. Altogether, these findings support an essential role for the *YWHAZ* gene in brain function and development, and together with our current report, highlight its contribution to neurodevelopmental disorders.

Given the functional evidence of this truncating mutation in *YWHAZ*, and several studies that consistently relate the 14-3-3 gene family to behavioural deficits in animal models, we further investigated the possible contribution of common and rare variants in this gene family to psychiatric disorders. The association study in our ASD sample failed to identify SNPs associated with the disorder, although the sample size was limited considering the small effect sizes typical of common variants in psychiatric disorders [13]. However, a previous study with similar sample size of adult ADHD patients identified a significant epistatic effect between *YWHAE* and two other members of this gene family, *YWHAZ* and *YWHAQ* [39]. When extending our analyses into larger samples of the PGC, BroadABC and iPSYCH, we found that common variants were gene-based associated with several psychiatric phenotypes in four of the seven 14-3-3 genes (*SFN*, *YWHAB*, *YWHAE* and *YWHAZ*), although only the association between *YWHAE* and schizophrenia survived correction for multiple testing. In line with this result, several previous studies found an association of genetic variants in *YWHAE* with schizophrenia [34,55].

The possible implication of 14-3-3 members in schizophrenia was also suggested in animal studies, which showed that 14-3-3ε deficient mice present alterations in hippocampal and cortical structures

due to defects in neurogenesis and neuronal migration [92,93]. In addition, 14-3-3ε deficient mice exhibit behavioural phenotypes, such as increased motor activity and decreased working memory and are used as schizophrenia-related models [34,55]. Furthermore, 14-3-3-mediated signaling seems to be strongly affected in schizophrenia patients [47,48], supporting the association of this protein family with the disorder. Interestingly, a polymorphism (rs28365859) in *YWHAE* associated with schizophrenia [34] correlates with differences in volumes of different brain regions in patients [94,95]. Taken together, these data support the contribution of *YWHAE* to schizophrenia and general brain development.

We also investigated the impact of rare variants in a small European sample of ASD sequencing all seven 14-3-3 genes. We identified 9 rare variants in 3 genes (*YWHAE*, *YWHAB* and *SFN*), including two rare variants in *SFN* that were predicted to be pathogenic (p.E75del and p.T165S) and were present on the same chromosome of the same patient. However, the functional characterization of the effect of these latter variants showed no effect on dimerization, although they may have other functional consequences.

We then expanded the analysis of rare variants to larger samples using available sequencing datasets of ASD and schizophrenia, which pinpointed a plausible impact for rare variants in ASD when we combined data of all 14-3-3 genes. Interestingly, the effect of this association was driven by URVs from *YWHAE* and *YWHAZ*. Indeed, several exome sequencing studies in ASD patients found rare variants in *YWHAZ* and *YWHAG* [19,96,97]. We also explored sequencing data from the ASC and SCHEMA consortia for each of the 14-3-3 genes individually and an enrichment was observed only for SNVs in *YWHAZ* in the schizophrenia dataset.

Interestingly, both the expression and the splicing of several 14-3-3 genes are regulated by RNA-binding proteins that are relevant to ASD and other psychiatric disorders. In particular, RBFOX1 (RNA binding fox-1 homolog 1) regulates *YWHAE*, *YWHAG*, *YWHAQ* and *YWHAZ*, whereas FMRP (Fragile X mental retardation protein) binds *YWHAG* [98,99]. Interestingly, *RBFOX1* has been reported in several ASD genetic studies [97,100–103], as well as in studies of aggressive behaviour and several psychiatric disorders [74,104–106]. Also, several targets of FMRP have been suggested to play a role in ASD [98,107,108].

Thus, our results indicate that combined rare variants in 14-3-3 genes may contribute to ASD, and that common variants from 14-3-3 family members significantly associate with schizophrenia. Also, when 14-3-3 genes are considered individually, rare variants in the *YWHAZ* gene are shown to contribute to schizophrenia.

Finally, we have systematically gathered and analysed previous transcriptomic data that suggest significant alteration of the expression of 14-3-3 genes in ASD and schizophrenia in several brain regions. Altered levels of the 14-3-3 family have previously been reported in ASD and in schizophrenia patients: 14-3-3 protein levels are diminished in platelets and pineal glands in ASD patients [29,44] and 14-3-3ζ levels are reduced in post-mortem brains of schizophrenia patients [109]. Moreover, an increased expression of *SFN* and decreased expression of *YWHAB*, *YWHAE*, *YWHAG* and *YWHAQ* mRNA have been reported in leukocytes of schizophrenia patients [45]. Our analysis of available transcriptomic data is in line with these results, as we found an altered brain expression of six of the seven 14-3-3 genes in ASD and schizophrenia patients: a robust increased expression of *SFN* in both ASD and schizophrenia (reflecting the top 9th and 36th differentially expressed gene in the transcriptome, respectively [110]; and a decreased expression of the other five 14-3-3 genes).

**5. Conclusions**

Our work combines functional studies, association studies, sequencing of a European sample, and interrogation of available genetic datasets that implicate the 14-3-3 gene family in ASD and schizophrenia, suggesting shared genetics between these two disorders.

## References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*; American Psychiatric Association: Washington, DC, USA, 2013.
2. Toma, C. Genetic variation across phenotypic severity of autism. *Trends Genet.* **2020**, *36*, 228–231. [CrossRef]
3. Hodges, H.; Fealko, C.; Soares, N. Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. *Transl. Pediatr.* **2020**, *9*, S55–S65. [CrossRef] [PubMed]
4. Christensen, D.L.; Maenner, M.J.; Bilder, D.; Constantino, J.N.; Daniels, J.; Durkin, M.S.; Fitzgerald, R.T.; Kurzius-Spencer, M.; Pettygrove, S.D.; Robinson, C.; et al. Prevalence and characteristics of Autism spectrum disorder among children aged 4 years-early autism and developmental disabilities monitoring network, seven sites, United States, 2010, 2012, and 2014. *MMWR Surveill. Summ.* **2019**, *68*, 1–19. [CrossRef]
5. Ronald, A.; Hoekstra, R.A. Autism spectrum disorders and autistic traits: A decade of new twin studies. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2011**, *156*, 255–274. [CrossRef] [PubMed]
6. Tick, B.; Bolton, P.; Happé, F.; Rutter, M.; Rijsdijk, F. Heritability of autism spectrum disorders: A meta-analysis of twin studies. *J. Child Psychol. Psychiatry Allied Discip.* **2016**, *57*, 585–595. [CrossRef] [PubMed]

7.  Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Hultman, C.; Larsson, H.; Reichenberg, A. The heritability of Autism spectrum disorder analysis method B. *JAMA* **2017**, *318*, 1182–1184. [CrossRef]

8.  Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat. Rev. Neurosci.* **2015**, *16*, 551–563. [CrossRef]

9.  De Rubeis, S.; Buxbaum, J.D. Recent advances in the genetics of Autism spectrum disorder. *Curr. Neurol. Neurosci. Rep.* **2015**, *15*, 36. [CrossRef]

10. Wang, K.; Zhang, H.; Ma, D.; Bucan, M.; Glessner, J.T.; Abrahams, B.S.; Salyakina, D.; Imielinski, M.; Bradfield, J.P.; Sleiman, P.M.A.; et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **2009**, *459*, 528–533. [CrossRef]

11. Anney, R.; Klei, L.; Pinto, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; Sykes, N.; Pagnamenta, A.T.; et al. A genome-wide scan for common alleles affecting risk for autism. *Hum. Mol. Genet.* **2010**, *19*, 4072–4082. [CrossRef]

12. Weiss, L.A.; Arking, D.E.; Daly, M.J.; Chakravarti, A.; Brune, C.W.; West, K.; O'Connor, A.; Hilton, G.; Tomlinson, R.L.; West, A.B.; et al. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **2009**, *461*, 802–808. [CrossRef] [PubMed]

13. Torrico, B.; Chiocchetti, A.G.; Bacchelli, E.; Trabetti, E.; Hervás, A.; Franke, B.; Buitelaar, J.K.; Rommelse, N.; Yousaf, A.; Duketis, E.; et al. Lack of replication of previous autism spectrum disorder GWAS hits in European populations. *Autism Res.* **2017**, *10*, 202–211. [CrossRef] [PubMed]

14. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444. [CrossRef]

15. Geisheker, M.R.; Heymann, G.; Wang, T.; Coe, B.P.; Turner, T.N.; Stessman, H.A.F.; Hoekzema, K.; Kvarnung, M.; Shaw, M.; Friend, K.; et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat. Neurosci.* **2017**, *20*, 1043–1051. [CrossRef] [PubMed]

16. Iossifov, I.; O'Roak, B.J.; Sanders, S.J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H.A.; Witherspoon, K.T.; Vives, L.; Patterson, K.E.; et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **2014**, *515*, 216–221. [CrossRef]

17. Roak, O.B.J.; Vives, L.; Girirajan, S.; Karakoc, E.; Krumm, N.; Coe, B.P.; Levy, R.; Ko, A.; Lee, C.; Smith, J.D.; et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **2012**, *485*, 246–250. [CrossRef]

18. Torrico, B.; Shaw, A.D.; Mosca, R.; Vivó-Luque, N.; Hervás, A.; Fernàndez-Castillo, N.; Aloy, P.; Bayés, M.; Fullerton, J.M.; Cormand, B.; et al. Truncating variant burden in high-functioning autism and pleiotropic effects of LRP1 across psychiatric phenotypes. *J. Psychiatry Neurosci.* **2019**, *44*, 350–359. [CrossRef]

19. Satterstrom, F.K.; Kosmicki, J.A.; Wang, J.; Breen, M.S.; De Rubeis, S.; An, J.-Y.; Peng, M.; Collins, R.; Grove, J.; Klei, L.; et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* **2020**, *180*, 568–584. [CrossRef]

20. Bacchelli, E.; Loi, E.; Cameli, C.; Moi, L.; Vega Benedetti, A.; Blois, S.; Fadda, A.; Bonora, E.; Mattu, S.; Fadda, R.; et al. Analysis of a sardinian multiplex family with autism spectrum disorder points to post-synaptic density gene variants and identifies CAPG as a functionally relevant candidate gene. *J. Clin. Med.* **2019**, *8*, 212. [CrossRef]

21. Chapman, N.H.; Nato, A.Q.; Bernier, R.; Ankenman, K.; Sohi, H.; Munson, J.; Patowary, A.; Archer, M.; Blue, E.M.; Webb, S.J.; et al. Whole exome sequencing in extended families with autism spectrum disorder implicates four candidate genes. *Hum. Genet.* **2015**, *134*, 1055–1068. [CrossRef]

22. Toma, C.; Torrico, B.; Hervás, A.; Valdés-Mas, R.; Tristán-Noguero, A.; Padillo, V.; Maristany, M.; Salgado, M.; Arenas, C.; Puente, X.S.; et al. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. *Mol. Psychiatry* **2014**, *19*, 784–790. [CrossRef] [PubMed]

23. Aghazadeh, Y.; Papadopoulos, V. The role of the 14-3-3 protein family in health, disease, and drug development. *Drug Discov. Today* **2016**, *21*, 278–287. [CrossRef] [PubMed]

24. Cornell, B.; Toyo-oka, K. 14-3-3 proteins in brain development: Neurogenesis, neuronal migration and neuromorphogenesis. *Front. Mol. Neurosci.* **2017**, *10*, 1–17. [CrossRef] [PubMed]

25. Hermeking, H.; Benzinger, A. 14-3-3 Proteins in cell cycle regulation. *Semin. Cancer Biol.* **2006**, *16*, 183–192. [CrossRef] [PubMed]

26. Aitken, A. 14-3-3 proteins: A historic overview. *Semin. Cancer Biol.* **2006**, *16*, 162–172. [CrossRef]

27. Kaplan, A.; Kent, C.B.; Charron, F.; Fournier, A.E. Switching responses: Spatial and temporal regulators of axon guidance. *Mol. Neurobiol.* **2014**, *49*, 1077–1086. [CrossRef]
28. Toma, C.; Rossi, M.; Sousa, I.; Blasi, F.; Bacchelli, E.; Alen, R.; Vanhala, R.; Monaco, A.P.; Järvelä, I.; Maestrini, E. Is ASMT a susceptibility gene for autism spectrum disorders? A replication study in European populations. *Mol. Psychiatry* **2007**, *12*, 977–979. [CrossRef]
29. Pagan, C.; Goubran-Botros, H.; Delorme, R.; Benabou, M.; Lemière, N.; Murray, K.; Amsellem, F.; Callebert, J.; Chaste, P.; Jamain, S.; et al. Disruption of melatonin synthesis is associated with impaired 14-3-3 and miR-451 levels in patients with autism spectrum disorders. *Sci. Rep.* **2017**, *7*, 2096. [CrossRef]
30. Melke, J.; Goubran Botros, H.; Chaste, P.; Betancur, C.; Nygren, G.; Anckarsäter, H.; Rastam, M.; Ståhlberg, O.; Gillberg, I.C.; Delorme, R.; et al. Abnormal melatonin synthesis in autism spectrum disorders. *Mol. Psychiatry* **2008**, *13*, 90–98. [CrossRef]
31. Toyo-Oka, K.; Wachi, T.; Hunt, R.F.; Baraban, S.C.; Taya, S.; Ramshaw, H.; Kaibuchi, K.; Schwarz, Q.P.; Lopez, A.F.; Wynshaw-Boris, A. 14-3-3E and Z regulate neurogenesis and differentiation of neuronal progenitor cells in the developing brain. *J. Neurosci.* **2014**, *34*, 12168–12181. [CrossRef]
32. Xu, X.; Jaehne, E.J.; Greenberg, Z.; McCarthy, P.; Saleh, E.; Parish, C.L.; Camera, D.; Heng, J.; Haas, M.; Baune, B.T.; et al. 14-3-3ζ deficient mice in the BALB/c background display behavioural and anatomical defects associated with neurodevelopmental disorders. *Sci. Rep.* **2015**, *5*. [CrossRef] [PubMed]
33. Cheah, P.S.; Ramshaw, H.S.; Thomas, P.Q.; Toyo-Oka, K.; Xu, X.; Martin, S.; Coyle, P.; Guthridge, M.A.; Stomski, F.; Van Den Buuse, M.; et al. Neurodevelopmental and neuropsychiatric behaviour defects arise from 14-3-3ζ deficiency. *Mol. Psychiatry* **2012**, *17*, 451–466. [CrossRef] [PubMed]
34. Ikeda, M.; Hikita, T.; Taya, S.; Uraguchi-asaki, J.; Toyo-Oka, K.; Wynshaw-boris, A.; Ujike, H.; Inada, T.; Takao, K.; Miyakawa, T.T.; et al. Identification of YWHAE, a gene encoding 14-3-3epsilon, as a possible susceptibility gene for schizophrenia. *Hum. Mol. Genet.* **2008**, *17*, 3212–3222. [CrossRef] [PubMed]
35. Jia, Y.; Yu, X.; Zhang, B.; Yuan, Y.; Xu, Q.; Shen, Y.; Shen, Y. An association study between polymorphisms in three genes of 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein) family and paranoid schizophrenia in northern Chinese population. *Eur. Psychiatry* **2004**, *19*, 377–379. [CrossRef]
36. Li, Z.; Chen, J.; Yu, H.; He, L.; Xu, Y.; Zhang, D.; Yi, Q.; Li, C.; Li, X.; Shen, J.; et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* **2017**, *49*, 1576–1583. [CrossRef] [PubMed]
37. Oldmeadow, C.; Mossman, D.; Evans, T.-J.; Holliday, E.G.; Tooney, P.A.; Cairns, M.J.; Wu, J.; Carr, V.; Attia, J.R.; Scott, R.J. Combined analysis of exon splicing and genome wide polymorphism data predict schizophrenia risk loci. *J. Psychiatr. Res.* **2014**, *52*, 44–49. [CrossRef]
38. Wong, A.H.C.; Macciardi, F.; Klempan, T.; Kawczynski, W.; Barr, C.L.; Lakatoo, S.; Wong, M.; Buckle, C.; Trakalo, J.; Boffa, E.; et al. Identification of candidate genes for psychosis in rat models, and possible association between schizophrenia and the 14-3-3η gene. *Mol. Psychiatry* **2003**, *8*, 156–166. [CrossRef] [PubMed]
39. Jacobsen, K.K.; Kleppe, R.; Johansson, S.; Zayats, T.; Haavik, J. Epistatic and gene wide effects in YWHA and aromatic amino hydroxylase genes across ADHD and other common neuropsychiatric disorders: Association with YWHAE. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2015**, *168*, 423–432. [CrossRef]
40. Grover, D.; Verma, R.; Goes, F.S.; Belmonte-Mahon, P.L.; Gershon, E.S.; McMahon, F.J.; Potash, J.B.; Steele, J.; Kassem, L.; Lopez, V.; et al. Family-based association of YWHAH in psychotic bipolar disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2009**, *150*, 977–983. [CrossRef]
41. Liu, J.; Zhang, H.X.; Li, Z.Q.; Li, T.; Li, J.Y.; Wang, T.; Li, Y.; Feng, G.Y.; Shi, Y.Y.; He, L. The YWHAE gene confers risk to major depressive disorder in the male group of Chinese Han population. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* **2017**, *77*, 172–177. [CrossRef]
42. Yanagi, M.; Shirakawa, O.; Kitamura, N.; Okamura, K.; Sakurai, K.; Nishiguchi, N.; Hashimoto, T.; Nushida, H.; Ueno, Y.; Kanbe, D.; et al. Association of 14-3-3 ε gene haplotype with completed suicide in Japanese. *J. Hum. Genet.* **2005**, *50*, 210–216. [CrossRef] [PubMed]
43. Garbett, K.; Ebert, P.J.; Mitchell, A.; Lintas, C.; Manzi, B.; Mirnics, K.; Persico, A.M. Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiol. Dis.* **2008**, *30*, 303–311. [CrossRef] [PubMed]

44. Pagan, C.; Delorme, R.; Callebert, J.; Goubran-Botros, H.; Amsellem, F.; Drouot, X.; Boudebesse, C.; Dudal, K.L.; Ngo-Nguyen, N.; Laouamri, H.; et al. The serotonin-N-acetylserotonin-melatonin pathway as a biomarker for autism spectrum disorders. *Transl. Psychiatry* **2014**, *4*, e479. [CrossRef] [PubMed]

45. Qing, Y.; Sun, L.; Yang, C.; Jiang, J.; Yang, X.; Hu, X.; Cui, D.; Xu, Y.; He, L.; Han, D.; et al. Dysregulated 14-3-3 family in peripheral blood leukocytes of patients with Schizophrenia. *Sci. Rep.* **2016**, *6*. [CrossRef]

46. Saia-Cereda, V.M.; Cassoli, J.S.; Schmitt, A.; Falkai, P.; Nascimento, J.M.; Martins-de-Souza, D. Proteomics of the corpus callosum unravel pivotal players in the dysfunction of cell signaling, structure, and myelination in schizophrenia brains. *Eur. Arch. Psychiatry Clin. Neurosci.* **2015**, *265*, 601–612. [CrossRef]

47. Saia-Cereda, V.M.; Cassoli, J.S.; Martins-de-Souza, D.; Nascimento, J.M. Psychiatric disorders biochemical pathways unraveled by human brain proteomics. *Eur. Arch. Psychiatry Clin. Neurosci.* **2017**, *267*, 3–17. [CrossRef]

48. Schubert, K.O.; Föcking, M.; Cotter, D.R. Proteomic pathway analysis of the hippocampus in schizophrenia and bipolar affective disorder implicates 14-3-3 signaling, aryl hydrocarbon receptor signaling, and glucose metabolism: Potential roles in GABAergic interneuron pathology. *Schizophr. Res.* **2015**, *167*, 64–72. [CrossRef]

49. Bruno, D.L.; Anderlid, B.M.; Lindstrand, A.; Van Ravenswaaij-Arts, C.; Ganesamoorthy, D.; Lundin, J.; Martin, C.L.; Douglas, J.; Nowak, C.; Adam, M.P.; et al. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J. Med. Genet.* **2010**, *47*, 299–311. [CrossRef]

50. Capra, V.; Mirabelli-Badenier, M.; Stagnaro, M.; Rossi, A.; Tassano, E.; Gimelli, S.; Gimelli, G. Identification of a rare 17p13.3 duplication including the BHLHA9 and YWHAE genes in a family with developmental delay and behavioural problems. *BMC Med. Genet.* **2012**, *13*, 93. [CrossRef]

51. Nagamani, S.C.S.; Zhang, F.; Shchelochkov, O.A.; Bi, W.; Ou, Z.; Scaglia, F.; Probst, F.J.; Shinawi, M.; Eng, C.; Hunter, J.V.; et al. Microdeletions including YWHAE in the Miller-Dieker syndrome region on chromosome 17p13.3 result in facial dysmorphisms, growth restriction, and cognitive impairment. *J. Med. Genet.* **2009**, *46*, 825–833. [CrossRef]

52. Blazejewski, S.M.; Bennison, S.A.; Smith, T.H.; Toyo-oka, K. Neurodevelopmental genetic diseases associated with microdeletions and microduplications of chromosome 17p13.3. *Front. Genet.* **2018**, *9*, 9. [CrossRef] [PubMed]

53. Curry, C.J.; Rosenfeld, J.A.; Grant, E.; Gripp, K.W.; Anderson, C.; Aylsworth, A.S.; Saad, T.B.; Chizhikov, V.V.; Dybose, G.; Fagerberg, C.; et al. The duplication 17p13.3 phenotype: Analysis of 21 families delineates developmental, behavioral and brain abnormalities, and rare variant phenotypes. *Am. J. Med. Genet. Part A* **2013**, *161*, 1833–1852. [CrossRef] [PubMed]

54. Ramocki, M.B.; Bartnik, M.; Szafranski, P.; Kołodziejska, K.E.; Xia, Z.; Bravo, J.; Miller, G.S.; Rodriguez, D.L.; Williams, C.A.; Bader, P.I.; et al. Recurrent distal 7q11.23 deletion including HIP1 and YWHAG identified in patients with intellectual disabilities, epilepsy, and neurobehavioral problems. *Am. J. Hum. Genet.* **2010**, *87*, 857–865. [CrossRef] [PubMed]

55. Wachi, T.; Cornell, B.; Toyo-oka, K. Complete ablation of the 14-3-3epsilon protein results in multiple defects in neuropsychiatric behaviors. *Behav. Brain Res.* **2017**, *319*, 31–36. [CrossRef]

56. Kim, D.E.; Cho, C.H.; Sim, K.M.; Kwon, O.; Hwang, E.M.; Kim, H.W.; Park, J.Y. 14-3-3Γ haploinsufficient mice display hyperactive and stress-sensitive behaviors. *Exp. Neurobiol.* **2019**, *28*, 43–53. [CrossRef]

57. Graham, K.; Zhang, J.; Qiao, H.; Wu, Y.; Zhou, Y. Region-specific inhibition of 14-3-3 proteins induces psychomotor behaviors in mice. *NPJ Schizophr.* **2019**, *5*, 1. [CrossRef]

58. Ghorbani, S.; Fossbakk, A.; Jorge-Finnigan, A.; Flydal, M.I.; Haavik, J.; Kleppe, R. Regulation of tyrosine hydroxylase is preserved across different homo-and heterodimeric 14-3-3 proteins. *Amino Acids* **2016**, *48*, 1221–1229. [CrossRef]

59. Itagaki, C.; Isobe, T.; Taoka, M.; Natsume, T.; Nomura, N.; Horigome, T.; Omata, S.; Ichinose, H.; Nagatsu, T.; Greene, L.A.; et al. Stimulus-coupled interaction of tyrosine hydroxylase with 14-3-3 proteins. *Biochemistry* **1999**, *38*, 15673–15680. [CrossRef]

60. Kleppe, R.; Rosati, S.; Jorge-Finnigan, A.; Alvira, S.; Ghorbani, S.; Haavik, J.; Valpuesta, J.M.; Heck, A.J.R.; Martinez, A. Phosphorylation dependence and stoichiometry of the complex formed by tyrosine hydroxylase and 14-3-3γ. *Mol. Cell. Proteom.* **2014**, *13*, 2017–2030. [CrossRef]

61. Moreno, E.; Vaz, S.H.; Cai, N.S.; Ferrada, C.; Quiroz, C.; Barodia, S.K.; Kabbani, N.; Canela, E.I.; McCormick, P.J.; Lluis, C.; et al. Dopamine-galanin receptor heteromers modulate cholinergic neurotransmission in the rat ventral hippocampus. *J. Neurosci.* **2011**, *31*, 7412–7423. [CrossRef]

62. Rutter, M.; LeCouteur, A.; Lord, C. *Autism Diagnostic Interview-Revised (ADI-R)*; Western Psychological Services: Torrance, CA, USA, 2003.

63. Lord, C.; Rutter, M.; DiLavore, P.; Risi, S.; Gotham, K.; Bishop, L.; RJ, L.; Guthrie, W. *Autism Diagnostic Observation Schedule*, 2nd ed.; Western Psychological Services: Torrance, CA, USA, 2012.

64. Miller, S.A.; Dykes, D.D.; Polesky, H.F. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* **1988**, *6*, 1215. [CrossRef] [PubMed]

65. Barrett, J.C.; Fry, B.; Maller, J.; Daly, M.J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **2005**, *21*, 263–265. [CrossRef] [PubMed]

66. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef] [PubMed]

67. Demontis, D.; Walters, R.K.; Martin, J.; Mattheisen, M.; Als, T.D.; Agerbo, E.; Baldursson, G.; Belliveau, R.; Bybjerg-Grauholm, J.; Bækvad-Hansen, M.; et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **2019**, *51*, 63–75. [CrossRef] [PubMed]

68. Tielbeek, J.J.; Johansson, A.; Polderman, T.J.C.; Rautiainen, M.R.; Jansen, P.; Taylor, M.; Tong, X.; Lu, Q.; Burt, A.S.; Tiemeier, H.; et al. Genome-wide association studies of a broad spectrum of antisocial behavior. *JAMA Psychiatry* **2017**, *74*, 1242–1250. [CrossRef] [PubMed]

69. Meier, S.M.; Trontti, K.; Purves, K.L.; Als, T.D.; Grove, J.; Laine, M.; Pedersen, M.G.; Bybjerg-Grauholm, J.; Bækved-Hansen, M.; Sokolowska, E.; et al. Genetic variants associated with anxiety and stress-related disorders: A genome-wide association study and mouse-model study. *JAMA Psychiatry* **2019**, *76*, 924–932. [CrossRef]

70. Stahl, E.A.; Breen, G.; Forstner, A.J.; McQuillin, A.; Ripke, S.; Trubetskoy, V.; Mattheisen, M.; Wang, Y.; Coleman, J.R.I.; Gaspar, H.A.; et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **2019**, *51*, 793–803. [CrossRef]

71. Wray, N.R.; Ripke, S.; Mattheisen, M.; Trzaskowski, M.; Byrne, E.M.; Abdellaoui, A.; Adams, M.J.; Agerbo, E.; Air, T.M.; Andlauer, T.M.F.; et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **2018**, *50*, 668–681. [CrossRef]

72. Arnold, P.D.; Askland, K.D.; Barlassina, C.; Bellodi, L.; Bienvenu, O.J.; Black, D.; Bloch, M.; Brentani, H.; Burton, C.L.; Camarena, B.; et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **2018**, *23*, 1181–1188. [CrossRef]

73. Ripke, S.; Neale, B.M.; Corvin, A.; Walters, J.T.R.; Farh, K.H.; Holmans, P.A.; Lee, P.; Bulik-Sullivan, B.; Collier, D.A.; Huang, H.; et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **2014**, *511*, 421–427. [CrossRef]

74. Cross-Disorder Group of the Psychiatric Genomics Consortium. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* **2019**, *179*, 1469–1482. [CrossRef] [PubMed]

75. De Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **2015**, *11*, 4. [CrossRef] [PubMed]

76. Kurosaki, T.; Maquat, L.E. Nonsense-mediated mRNA decay in humans at a glance. *J. Cell Sci.* **2016**, *129*, 461–467. [CrossRef] [PubMed]

77. Khajavi, M.; Inoue, K.; Lupski, J.R. Nonsense-mediated mRNA decay modulates clinical outcome of genetic disease. *Eur. J. Hum. Genet.* **2006**, *14*, 1074–1081. [CrossRef] [PubMed]

78. Obsil, T.; Obsilova, V. Structural basis of 14-3-3 protein functions. *Semin. Cell Dev. Biol.* **2011**, *22*, 663–672. [CrossRef]

79. Halskau, Ø.J.; Ying, M.; Baumann, A.; Kleppe, R.; Rodriguez-Larrea, D.; Almås, B.; Haavik, J.; Martinez, A. Three-way interaction between 14-3-3 proteins, the N-terminal region of tyrosine hydroxylase, and negatively charged membranes. *J. Biol. Chem.* **2009**, *284*, 32758–32769. [CrossRef]

80. Uhart, M.; Bustos, D.M. Human 14-3-3 paralogs differences uncovered by cross-talk of phosphorylation and lysine acetylation. *PLoS ONE* **2013**, *8*, 1–16. [CrossRef]

81. Rosenquist, M.; Sehnke, P.; Ferl, R.J.; Sommarin, M.; Larsson, C. Evolution of the 14-3-3 protein family: Does the large number of isoforms in multicellular organisms reflect functional specificity? *J. Mol. Evol.* **2000**, *51*, 446–458. [CrossRef]

82. Berg, D.; Holzmann, C.; Riess, O. 14-3-3 Proteins in the nervous system. *Nat. Rev. Neurosci.* **2003**, *4*, 752–762. [CrossRef]

83. Kleppe, R.; Martinez, A.; Døskeland, S.O.; Haavik, J. The 14-3-3 proteins in regulation of cellular metabolism. *Semin. Cell Dev. Biol.* **2011**, *22*, 713–719. [CrossRef]

84. Cabana-Domínguez, J.; Torrico, B.; Reif, A.; Fernàndez-Castillo, N.; Cormand, B. Exploring the genetic contribution of the dopaminergic and serotonergic pathways in psychiatric disorders. *Transl. Psychiatry* **2019**, *9*, 242. [CrossRef]

85. Ribasés, M.; Ramos-Quiroga, J.A.; Hervás, A.; Sánchez-Mora, C.; Bosch, R.; Bielsa, A.; Gastaminza, X.; Lesch, K.P.; Reif, A.; Renner, T.J.; et al. Candidate system analysis in ADHD: Evaluation of nine genes involved in dopaminergic neurotransmission identifies association with DRD1. *World J. Biol. Psychiatry* **2012**, *13*, 281–292. [CrossRef] [PubMed]

86. Toma, C.; Hervás, A.; Balmaña, N.; Salgado, M.; Maristany, M.; Vilella, E.; Aguilera, F.; Orejuela, C.; Cuscó, I.; Gallastegui, F.; et al. Neurotransmitter systems and neurotrophic factors in autism: Association study of 37 genes suggests involvement of DDC. *World J. Biol. Psychiatry* **2013**, *14*, 516–527. [CrossRef] [PubMed]

87. Treister, R.; Pud, D.; Ebstein, R.P.; Laiba, E.; Gershon, E.; Haddad, M.; Eisenberg, E. Associations between polymorphisms in dopamine neurotransmitter pathway genes and pain response in healthy humans. *Pain* **2009**, *147*, 187–193. [CrossRef]

88. Wood, P.B.; Holman, A.J. An elephant among us: The role of dopamine in the pathophysiology of fibromyalgia. *J. Rheumatol.* **2009**, *36*, 221–224. [CrossRef]

89. Ghorbani, S.; Szigetvari, P.D.; Haavik, J.; Kleppe, R. Serine 19 phosphorylation and 14-3-3 binding regulate phosphorylation and dephosphorylation of tyrosine hydroxylase on serine 31 and serine 40. *J. Neurochem.* **2020**, *152*, 29–47. [CrossRef]

90. Ramshaw, H.; Xu, X.; Jaehne, E.J.; McCarthy, P.; Greenberg, Z.; Saleh, E.; McClure, B.; Woodcock, J.; Kabbara, S.; Wiszniak, S.; et al. Locomotor hyperactivity in 14-3-3ζ KO mice is associated with dopamine transporter dysfunction. *Transl. Psychiatry* **2013**, *3*, e327. [CrossRef]

91. Qiao, H.; Foote, M.; Graham, K.; Wu, Y.; Zhou, Y. 14-3-3 proteins are required for hippocampal long-term potentiation and associative learning and memory. *J. Neurosci.* **2014**, *34*, 4801–4808. [CrossRef]

92. Toyo-Oka, K.; Shionoya, A.; Gambello, M.J.; Cardoso, C.; Leventer, R.; Ward, H.L.; Ayala, R.; Tsai, L.H.; Dobyns, W.; Ledbetter, D.; et al. 14-3-3ε is important for neuronal migration by binding to NUDEL: A molecular explanation for Miller-Dieker syndrome. *Nat. Genet.* **2003**, *34*, 274–285. [CrossRef]

93. Cornell, B.; Wachi, T.; Zhukarev, V.; Toyo-oka, K. Regulation of neuronal morphogenesis by 14-3-3epsilon (Ywhae) via the microtubule binding protein, doublecortin. *Hum. Mol. Genet.* **2016**, *25*, 4405–4418. [CrossRef]

94. Takahashi, T.; Nakamura, Y.; Nakamura, Y.; Aleksic, B.; Takayanagi, Y.; Furuichi, A.; Kido, M.; Nakamura, M.; Sasabayashi, D.; Ikeda, M.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and orbitofrontal sulcogyral pattern in patients with schizophrenia and healthy subjects. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* **2014**, *51*, 166–171. [CrossRef] [PubMed]

95. Kido, M.; Nakamura, Y.; Nemoto, K.; Takahashi, T.; Aleksic, B.; Furuichi, A.; Nakamura, Y.; Ikeda, M.; Noguchi, K.; Kaibuchi, K.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and brain morphology in Schizophrenia: A voxel-based morphometric study. *PLoS ONE* **2014**, *9*, E103571. [CrossRef] [PubMed]

96. Lim, E.T.; Uddin, M.; De Rubeis, S.; Chan, Y.; Kamumbu, A.S.; Zhang, X.; D'Gama, A.M.; Kim, S.N.; Hill, R.S.; Goldberg, A.P.; et al. Rates, distribution and implications of postzygotic mosaic mutations in Autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*, 1217–1224. [CrossRef]

97. De Rubeis, S.; He, X.; Goldberg, A.P.; Poultney, C.S.; Samocha, K.; Cicek, A.E.; Kou, Y.; Liu, L.; Fromer, M.; Walker, S.; et al. Synaptic, transcriptional and chromatin genes disrupted in Autism. *Nature* **2014**, *515*, 209–215. [CrossRef] [PubMed]

98. Darnell, J.C.; Van Driesche, S.J.; Zhang, C.; Hung, K.Y.S.; Mele, A.; Fraser, C.E.; Stone, E.F.; Chen, C.; Fak, J.J.; Chi, S.W.; et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and Autism. *Cell* **2011**, *146*, 247–261. [CrossRef]

99. Lee, J.A.; Damianov, A.; Lin, C.H.; Fontes, M.; Parikshak, N.N.; Anderson, E.S.; Geschwind, D.H.; Black, D.L.; Martin, K.C. Cytoplasmic Rbfox1 regulates the expression of synaptic and autism-related genes. *Neuron* **2016**, *89*, 113–128. [CrossRef]

100. Hamada, N.; Ito, H.; Nishijo, T.; Iwamoto, I.; Morishita, R.; Tabata, H.; Momiyama, T.; Nagata, K.I. Essential role of the nuclear isoform of RBFOX1, a candidate gene for Autism spectrum disorders, in the brain development. *Sci. Rep.* **2016**, *6*, 30805. [CrossRef]

101. Turner, T.N.; Hormozdiari, F.; Duyzend, M.H.; McClymont, S.A.; Hook, P.W.; Iossifov, I.; Raja, A.; Baker, C.; Hoekzema, K.; Stessman, H.A.; et al. Genome sequencing of Autism-affected families reveals disruption of putative noncoding regulatory DNA. *Am. J. Hum. Genet.* **2016**, *98*, 58–74. [CrossRef]

102. Voineagu, I.; Wang, X.; Johnston, P.; Lowe, J.K.; Tian, Y.; Horvath, S.; Mill, J.; Cantor, R.M.; Blencowe, B.J.; Geschwind, D.H. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **2011**, *474*, 380–386. [CrossRef]

103. Kanduri, C.; Kantojärvi, K.; Salo, P.M.; Vanhala, R.; Buck, G.; Blancher, C.; Lähdesmäki, H.; Järvelä, I. The landscape of copy number variations in Finnish families with autism spectrum disorders. *Autism Res.* **2016**, *9*, 9–16. [CrossRef]

104. Fernàndez-Castillo, N.; Gan, G.; van Donkelaar, M.M.J.; Vaht, M.; Weber, H.; Retz, W.; Meyer-Lindenberg, A.; Franke, B.; Harro, J.; Reif, A.; et al. RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur. Neuropsychopharmacol.* **2020**, *30*, 44–55. [CrossRef] [PubMed]

105. Lal, D.; Pernhorst, K.; Klein, K.M.; Reif, P.; Tozzi, R.; Toliat, M.R.; Winterer, G.; Neubauer, B.; Nürnberg, P.; Rosenow, F.; et al. Extending the phenotypic spectrum of RBFOX1 deletions: Sporadic focal epilepsy. *Epilepsia* **2015**, *56*, e129–e133. [CrossRef] [PubMed]

106. Wen, M.; Yan, Y.; Yan, N.; Chen, X.S.; Liu, S.Y.; Feng, Z.H. Upregulation of RBFOX1 in the malformed cortex of patients with intractable epilepsy and in cultured rat neurons. *Int. J. Mol. Med.* **2015**, *35*, 597–606. [CrossRef] [PubMed]

107. Budimirovic, D.B.; Kaufmann, W.E. What can we learn about autism from studying fragile X syndrome? *Dev. Neurosci.* **2011**, *33*, 379–394. [CrossRef]

108. Lai, A.; Valdez-Sinon, A.N.; Bassell, G.J. Regulation of RNA granules by FMRP and implications for neurological diseases. *Traffic* **2020**. [CrossRef]

109. English, J.A.; Pennington, K.; Dunn, M.J.; Cotter, D.R. The neuroproteomics of Schizophrenia. *Biol. Psychiatry* **2011**, *69*, 163–172. [CrossRef]

110. Gandal, M.J.; Zhang, P.; Hadjimichael, E.; Walker, R.L.; Chen, C.; Liu, S.; Won, H.; Van Bakel, H.; Varghese, M.; Wang, Y.; et al. Transcriptome-wide isoform-level dysregulation in ASD, Schizophrenia, and bipolar disorder. *Science* **2018**, *362*. [CrossRef]

# SUPPLEMENTARY FILES

**INVOLVEMENT OF THE 14-3-3 GENE FAMILY IN AUTISM SPECTRUM DISORDER AND**

**SCHIZOPHRENIA: GENETICS, TRANSCRIPTOMICS AND FUNCTIONAL ANALYSES**

Bàrbara Torrico, Ester Antón-Galindo, Noèlia Fernàndez-Castillo, Eva Rojo, Sadaf Ghorbani, Laura Pineda, Amaia Hervás, Isabel Rueda, Estefanía Moreno, Janice M. Fullerton, Vicent Casadó, Jan K. Buitelaar, Nanda Rommelse, Barbara Franke, Andreas Reif, Andreas G. Chiocchetti, Christine Freitag, Rune Kleppe, Jan Haavik, Claudio Toma, Bru Cormand

**This supplementary file includes:**

Additional methodological procedures

Tables S1 to S7

Figures S1 to S4

References

**ADDITIONAL METHODOLOGICAL PROCEDURES**

*DNA constructs and site-directed mutagenesis for expression in prokaryotes*

The expression vector pGEX-2TK (GE Healthcare, Little Chalfont, UK) was used to clone the cDNA of the wild-type (WT) *YWHAZ* human gene at 3' of its glutathione-s-transferase (GST) tag (GST-14-3-3ζ_WT). The mutated 14-3-3ζ form was obtained through a directed mutagenesis protocol using the *Site-Directed Mutagenesis kit* (Stratagene, La Jolla, CA, USA), inserting a thymine between the nucleotides 659 and 660 of the *YWHAZ* cDNA insert as per the observed p.L220Ffs*18 mutation (GST-14-3-3ζ_mut).

*Expression and purification in prokaryotes*

WT and mutated fusion proteins (GST-14-3-3ζ_WT and GST-14-3-3ζ_mut) were expressed in the *BL21 Codon Plus E. coli* strain (Stratagene) by a 4h induction at 30 ºC with 1 mM of 1-thio-β-D-galactopyranoxide (IPTG). Bacteria were lysed in a lysis buffer composed of phosphate buffered saline (PBS) 10 mM imidazole, 10 mM benzamidine, with 0.5 mg/mL lysozyme and protease inhibitor cocktail (Roche, Mannheim, Germany), using the French press. Then, the fusion proteins GST-14-3-3ζ_WT and GST-14-3-3ζ_mut were purified from the soluble fraction of the lysate through affinity chromatography using *Glutathione Sepharose 4B* (GE Healthcare). The homogeneity of the purified proteins was confirmed by SDS PAGE and quantified using theoretical absorbance at 280 nm (as the mutant has lost one Trp residue) as well as by protein staining of gels. Size exclusion chromatography was used to confirm dimeric state of soluble 14-3-3 proteins as described [1].

*Solubility test of 14-3-3ζ WT and mutated proteins*

Induction of the expression of the two fusion proteins with IPTG was performed at different temperatures and final induction times: 4 hours at 30 ºC, 5 hours at 25 ºC and 6 hours at 20 ºC. Cells were resuspended in the lysis buffer mentioned above and lysed by sonication. Two samples were obtained and kept in 1X sample buffer for SDS-PAGE: the total lysate sample (immediately after sonication) and the soluble sample, from the supernatant fraction after 10 min centrifugation at 13,000 *g*. Then, all samples were run on 12% denaturing polyacrylamide gels (SDS-PAGE) and patterns of WT and truncated proteins were compared.

*DNA constructs for expression in eukaryotes (BRET assays)*

Human cDNAs for WT or mutant *YWHAZ* were amplified from a carrier ASD proband, using sense and antisense primers harboring *Kpn*I and *Eco*RI restriction sites to clone the generated amplicons into a pcDNA3 vector (Promega, Madison, WI, USA). Subsequently, cDNAs for EYFP (enhanced yellow variant of GFP) and Rluc (*Renilla* luciferase) were amplified without their stop codons, adding *Kpn*I restriction sites at both ends of the amplicon, from pEYFP-N1 (Takara Bio Inc, Otsu, Shiga, Japan) and pRluc-N1 (PerkinElmer, Wellesley, MA, United States) vectors, respectively. The resulting fragments were subcloned in the previously generated pcDNA3_YWHAZ vectors at 5' of the *YWHAZ* insert and in-frame with its start codon. Human cDNA for *YWHAE*, cloned into pcDNA3.1 (Promega), was amplified without its stop codon using sense and antisense primers harboring *Bam*HI and *Eco*RI restriction sites. The resulting fragment

was subcloned at 5' of EYFP or Rluc to be in frame with their start codons in the pEYFP-N1 and pRluc-N1 vectors respectively.

Human cDNAs for WT and mutant *SFN* were obtained from the carrier ASD proband (MT_37.3), which contained the three described variants in the same chromosome. For that, cDNAs were amplified without including the stop codon, using primers harboring *Hind*III and *Bam*HI restriction sites. All the resulting fragments were then subcloned to be in frame with pEYFP or pRluc vectors.

Briefly, we obtained ten different plasmids: YFP-YWHAZ WT, YFP-YWHAZ mut, Rluc-YWHAZ WT, Rluc-YWHAZ mut, YWHAE-YFP, YWHAE-Rluc, SFN WT-YFP, SFN mut-YFP, SFN WT-Rluc, SFN mut-Rluc.


*CaCl$_2$ transfection of HEK293T cells for BRET assays*

HEK293T cells were cultured at 37ºC with 5% $CO_2$ with DMEM supplemented with 10% fetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin. For transfection, the day before cells were plated in 6-well plates at a density of $5\times10^5$ cells/well and media was replaced 2 hours before transfection. Individual transfection reactions were prepared in 1.5ml sterile tubes. The amount of plasmid(s) (see Supplementary Table 1) to be transfected was diluted in buffered water (HEPES 2.5mM, pH 7.3) and TE (tris, EDTA, pH 8, final concentration 0.03x) to obtain a final volume of 150µl in each sterile tube. Then, 14.9µl of CaCl$_2$ 2.5M was added to each tube while vortexing and 150µl of HBS 2x (280 mM NaCl, 50 mM HEPES, 40 mM Na$_2$HPO$_2$.2H$_2$O, pH 7.00) was subsequently added drop by drop while vortexing in order to mix well the solution. All the transfection solutions were kept at room temperature for 20 minutes and each solution was added to each well containing HEK293T. Media was replaced 17 hours after transfection.

**SUPPLEMENTARY TABLES**

**Table S1.** Acceptor and donor combinations of plasmids for BRET experiments.

| Donor \ Acceptor | YWHAE-YFP | SFN WT-YFP | SFN mut-YFP | DRD1-YFP |
|---|---|---|---|---|
| **Rluc-YWHAZ WT** | 250 / 5 - 1000 | - | - | 575 / 200 - 6500 |
| **Rluc-YWHAZ mut** | 2250 / 5 - 1500 | - | - | 4250 / 200 - 4000 |
| **YWHAE-Rluc** | - | - | - | - |
| **SFN WT-Rluc** | - | 2.5 / 10-180 | 2.5 / 10-180 | 9 / 100-3200 |
| **SFN mut-Rluc** | - | - | 5 / 10-180 | 12.5 / 200-2600 |
| **DRD1-Rluc** | 20 / 20 - 700 | 30 / 10 - 400 | 20 / 10 - 140 | - |

The amount of each plasmid is indicated as ng of donor / range of ng of acceptor. Donor amount is indicated as an average of the weight of plasmid used in the co-transfections in order to obtain a stable luminescence signal (around 150.000 bioluminescence units) at 10 min.

**Table S2.** European ASD samples used in the case-control association study and in the mutation screening.

| | Association study | | Mutation screening |
|---|---|---|---|
| | Cases (%M) | Controls (%M) | Cases (%M) |
| **Spanish** | 301 (87.3) | 300 (89.5) | 182 (86.8) |
| **Dutch** | 238 (78.6) | 235 (77.9) | 89 (79.5) |
| **German** | 188 (90.5) | 179 (83.2) | 14 (85.7) |
| **Total** | 727 (85) | 714 (84.7) | 285 (84.5) |

%M indicates the percentage of male individuals.

**Table S3.** Results from the ASD case-control association study (727 ASD cases and 714 controls) with tagSNPs across the 14-3-3 gene family in the overall sample under the additive model.

| Gene | Chr | tagSNP | Position GRCh37/hg19 | MAF (CEU) | Alleles | MAF in our sample | HWE (unaffected) | P-val (additive) |
|---|---|---|---|---|---|---|---|---|
| *SFN* | Chr1: 27189632-27190947 | rs200392321 + | 27187881 | 0.08 | T:G | - | - | - |
| | | rs1883660 | 27191411 | 0.128 | A:T | 0.1433 | 0.8864 | 0.024 |
| *YWHAQ* | Chr2: 9724095-9771184 | rs6734469 | 9735900 | 0.381 | A:G | 0.4523 | 0.6455 | 0.821 |
| | | rs13022460 | 9749492 | 0.275 | T:G | 0.2556 | 0.6179 | 0.707 |
| | | rs16867074 | 9750210 | 0.092 | G:C | 0.1082 | 0.6848 | 0.413 |
| | | rs17453675 | 9754407 | 0.242 | A:G | 0.1883 | 0.9014 | 0.965 |
| | | rs4668625 | 9765461 | 0.35 | A:G | 0.3243 | 0.7278 | 0.82 |
| | | rs13417081 | 9770050 | 0.295 | G:A | 0.368 | 0.935 | 0.893 |
| | | rs3762535 | 9773586 | 0.491 | C:T | 0.4598 | 0.2826 | 0.43 |
| | | rs4145375 | 9773678 | 0.094 | A:G | 0.0752 | 0.4093 | 0.782 |
| *YWHAG* | Chr7: 75956107-75988342 | rs2961037 | 75952212 | 0.5 | G:G | 0.494 | 0.447 | 0.797 |
| | | rs2961033 | 75964343 | 0.25 | T:C | 0.1912 | 0.2729 | 0.725 |
| | | rs10241401 # | 75975427 | 0.057 | G:A | 0.02954 | 1 | - |
| | | rs13247572 | 75983524 | 0.083 | G:A | 0.07367 | 0.763 | 0.311 |
| | | rs11765693 | 75985373 | 0.196 | A:G | 0.2811 | 0.4464 | 0.558 |
| *YWHAZ* | Chr8: 101930803-101965623 | rs17365305 # | 101933682 | 0.054 | G:A | 0.04128 | 0.6176 | - |
| | | rs4734497 | 101934971 | 0.36 | T:C | 0.289 | 0.3965 | 0.32 |
| | | rs17462921 | 101938901 | 0.176 | G:A | 0.1537 | 0.4824 | 0.287 |
| | | rs3134354 | 101948681 | 0.063 | C:G | 0.05836 | 1 | 0.604 |
| | | rs17366009 | 101957311 | 0.072 | T:C | 0.06192 | 0.5037 | 0.475 |
| | | rs3100052 | 101967139 | 0.396 | G:A | 0.3995 | 0.5808 | 0.694 |
| *YWHAB* | Chr20: 43514239-43537175 | rs6031849 | 43514337 | 0.188 | G:T | 0.3082 | 0.3325 | 0.327 |
| | | rs2425675 | 43534934 | 0.25 | G:A | 0.2897 | 0.1656 | 0.938 |
| | | rs6876 | 43535101 | 0.098 | T:C | 0.1322 | 0.5277 | 0.308 |
| | | rs2425678 | 43538152 | 0.438 | T:C | 0.402 | 0.1099 | 0.221 |
| *YWHAH* | Chr22: 32340478-32353590 | rs3827334 | 32338005 | 0.188 | A:C | 0.1601 | 0.4722 | 0.631 |
| | | rs929036 | 32339213 | 0.413 | C:T | 0.4477 | 0.5908 | 0.797 |
| | | rs2267172 | 32339782 | 0.064 | G:A | 0.07153 | 0.7718 | 1 |
| | | rs5998196 | 32355455 | 0.417 | T:C | 0.433 | 1 | 0.421 |
| *YWHAE* | Chr17:1247833-1303556 | rs11650689 | 1244992 | 0.348 | C:T | 0.36 | 0.6816 | 0.617 |
| | | rs9393 | 1248392 | 0.054 | A:G | 0.08007 | 1 | 0.316 |
| | | rs7224258 | 1255502 | 0.33 | G:C | 0.3085 | 0.6508 | 0.256 |
| | | rs7208041 | 1270562 | 0.241 | A:G | 0.2164 | 0.04387 | 0.891 |
| | | rs4790082 | 1278700 | 0.446 | G:A | 0.4872 | 0.5953 | 0.924 |
| | | rs17625475 | 1280109 | 0.107 | G:T | 0.09253 | 0.6231 | 0.188 |
| | | rs10521111 | 1281864 | 0.152 | A:G | 0.1357 | 0.03108 | 0.685 |
| | | rs16945811 | 1294614 | 0.078 | G:A | 0.07585 | 0.7944 | 0.377 |

Chr: chromosome; SNP: single nucleotide polimorphism; MAF: minor allele frequency in our sample; HWE: Hardy-Weinberg equilibrium p-value, calculated in our control sample; P-val: p-value for the association under the additive model. +, monomorphic SNP excluded from analyses; # SNPs with a MAF<0.05 excluded from analyses. Bonferroni threshold for multiple testing correction p=0.05/34 SNPs

**Table S4.** Description of the summary statistics of publicly available GWAS data of eight psychiatric disorders and the corresponding cross-disorder dataset used for gene-based and gene-set analyses.

| GWAS | Participants | Reference |
|---|---|---|
| Attention-Deficit Hyperactivity Disorder, ADHD | 19,099 Ca + 34,194 Co | Demontis et al. 2019 [2] |
| Antisocial Behavior, ASB | 16,400 | Tielbeek et al, 2017 [3] |
| Anxiety | 12,655 Ca + 19,255 Co | Meier et al. 2019 [4] |
| Autism Spectrum Disorder, ASD | 18,382 Ca + 27,969 Co | Grove et al. 2019 [5] |
| Bipolar Disorder, BD | 20,352 Ca + 31,358 Co | Stahl et al. 2019 [6] |
| Major Depressive Disorder, MDD | 59,851 Ca + 113,154 Co | Wray et al. 2018 [7] |
| Obsessive-Compulsive Disorder, OCD | 1,773 Ca + 6,122 Co + 915 trios | Arnold et al. 2018 [8] |
| Schizophrenia, SCZ | 32,405 Ca + 42,221 Co + 1,235 trios | Ripke et al. 2014 [9] |
| Cross-Disorder meta-analysis | 162,151 Ca + 276,846 Co | Lee et al. 2019 [10] |

Ca: Cases; Co: controls; GWAS, genome-wide association study.

**Table S5.** Experimental design of targeted next-generation sequencing: The coding regions of 14-3-3 genes was covered by 57 amplicons.

| Request_ID | Target_ID | Gene | Chr:start-end | N Amplicons | Total Bases | Covered Bases | Missed Bases | Coverage |
|---|---|---|---|---|---|---|---|---|
| IAD38961 | 503029 | *YWHAE* | chr17:1248735-98 | 1 | 64 | 64 | 0 | 1 |
| IAD38961 | 503026 | *YWHAE* | chr17:1257499-646 | 2 | 148 | 148 | 0 | 1 |
| IAD38961 | 503030 | *YWHAE* | chr17:1264380-597 | 3 | 218 | 218 | 0 | 1 |
| IAD38961 | 503028 | *YWHAE* | chr17:1265190-307 | 2 | 118 | 118 | 0 | 1 |
| IAD38961 | 503025 | *YWHAE* | chr17:1268147-357 | 2 | 211 | 211 | 0 | 1 |
| IAD38961 | 503027 | *YWHAE* | chr17:1303335-409 | 1 | 75 | 75 | 0 | 1 |
| IAD38961 | 484571 | *YWHAQ* | chr2:9725409-479 | 1 | 71 | 71 | 0 | 1 |
| IAD38961 | 484573 | *YWHAQ* | chr2:9727537-643 | 2 | 107 | 107 | 0 | 1 |
| IAD38961 | 484570 | *YWHAQ* | chr2:9728288-462 | 2 | 175 | 175 | 0 | 1 |
| IAD38961 | 484572 | *YWHAQ* | chr2:9731515-649 | 2 | 135 | 135 | 0 | 1 |
| IAD38961 | 484574 | *YWHAQ* | chr2:9770282-586 | 3 | 305 | 305 | 0 | 1 |
| IAD38961 | 503001 | *SFN* | chr1:27189698-190455 | 6 | 758 | 758 | 0 | 1 |

| IAD38961 | 484575 | *YWHAH* | chr22:32340714-811 | 1 | 98 | 72 | 26 | 0.735 |
|---|---|---|---|---|---|---|---|---|
| IAD38961 | 484576 | *YWHAH* | chr22:32352120-2784 | 5 | 665 | 665 | 0 | 1 |
| IAD38961 | 482745 | *YWHAB* | chr20:43530169-0479 | 3 | 311 | 311 | 0 | 1 |
| IAD38961 | 482743 | *YWHAB* | chr20:43532628-762 | 2 | 135 | 135 | 0 | 1 |
| IAD38961 | 482744 | *YWHAB* | chr20:43533603-777 | 2 | 175 | 175 | 0 | 1 |
| IAD38961 | 482742 | *YWHAB* | chr20:43534636-742 | 1 | 107 | 107 | 0 | 1 |
| IAD38961 | 482741 | *YWHAB* | chr20:43535017-084 | 1 | 68 | 68 | 0 | 1 |
| IAD38961 | 482387 | *YWHAG* | chr7:75958888-9555 | 5 | 668 | 668 | 0 | 1 |
| IAD38961 | 482388 | *YWHAG* | chr7:75988033-130 | 1 | 98 | 98 | 0 | 1 |
| IAD38961 | 482475 | *YWHAZ* | chr8:101932915-85 | 1 | 71 | 71 | 0 | 1 |
| IAD38961 | 482472 | *YWHAZ* | chr8:101936177-283 | 2 | 107 | 107 | 0 | 1 |
| IAD38961 | 482473 | *YWHAZ* | chr8:101936357-531 | 1 | 175 | 53 | 122 | 0.303 |
| IAD38961 | 482471 | *YWHAZ* | chr8:101937138-272 | 2 | 135 | 135 | 0 | 1 |
| IAD38961 | 482474 | *YWHAZ* | chr8:101960818-1122 | 3 | 305 | 305 | 0 | 1 |

For each coding exon the number of amplicons used to cover the exon length is reported. The overall coverage was 96.3% across the 14-3-3 gene family.

**Table S6.** Gene-set association results of the 14-3-3 family set of genes on eight different psychiatric phenotypes and in the cross-disorder meta-analysis dataset.

| | ADHD | ASB | Anxiety | ASD | BP | MDD | OCD | **SCZ** | Cross-disorder |
|---|---|---|---|---|---|---|---|---|---|
| *p value* | 0.251 | 0.476 | 0.270 | 0.835 | 0.110 | 0.772 | 0.090 | **0.015** | 0.112 |

*p*-values were calculated using MAGMA (v1.06) software. Nominal association is highlighted in bold. ADHD: attention-deficit hyperactivity disorder; ASB: antisocial behaviour; ASD: autism spectrum disorder; BP: bipolar disorder; MDD: major depressive disorder; OCD: obsessive-compulsive disorder; SCZ: schizophrenia.

**Table S7.** Rare variants identified in the seven 14-3-3 family genes in 285 European ASD patients.

| Individual | Sex | Comorbidity | Family type | Origin | Gene | Allelic change | Chr:position[a] | gene region | Protein | dbSNP | SIFT/ Provean[b] | PolyPhen2[c] | MAF (ExAC)[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MT_37.3 | M | - | S | mother | SFN | A/T | 1:27190196 | exon | p.T165S | rs77755255 | Damaging / Deleterious | Probably damaging | 2.47E-04 |
| | | | | mother | SFN | GGA/- | 1:27189925-27189927 | exon | p.E76del | - | - / Deleterious | - | 1.65E-05 |
| | | | | mother | SFN | C/T | 1:27190149 | exon | p.S149L | rs78707984 | Tolerated / Neutral | benign | 2.48E-04 |
| MT_11.3 | M | LD | S | unknown | SFN | C/T | 1:27189940 | exon | p.P79P | - | Tolerated / Neutral | - | - |
| MT_159.3 | M | - | S | unknown | YWHAE | G/A | 17:1248772 | exon; ncRNA | p.A246V | - | Tolerated / Neutral | benign | - |
| 10-09471 | - | - | | unknown | YWHAE | G/A | 17:1303445 | 5'UTR, ncRNA | - | rs139532375 | - | - | 3.56E-03 |
| SJD_64.3 | M | - | S | mother | YWHAB | A/C | 20:43530403 | exon | p.K77Q | rs142757633 | Damaging / Neutral | benign | 1.82E-03 |
| SJD_18.3 | M | - | | father | YWHAB | G/A | 20:43532595 | intronic | - | rs199806929 | - | - | 2.61E-03 |
| MT_160.3 | M | Epilepsy | Mx | mother | YWHAZ | ins-T | 8:101936203–204 | exon | p.L220F*fs*18 | - | - | - | - |

[a]GRCh37/hg19 assembly; [b]SIFT/Provean (Sorting Intolerant from Tolerant, http://sift.jcvi.org/); [c]PolyPhen2 (http://genetics.bwh.harvard.edu/pph2); [d]MAF (Minor Allele Frequency) of the variant in the ExAC Browser (http://exac.broadinstitute.org);). M: male; Mx, multiplex; LD: Language Delay; ncRNA: non-coding RNA; S, singleton; 5'UTR: 5' untranslated region.

**SUPPLEMENTARY FIGURES**
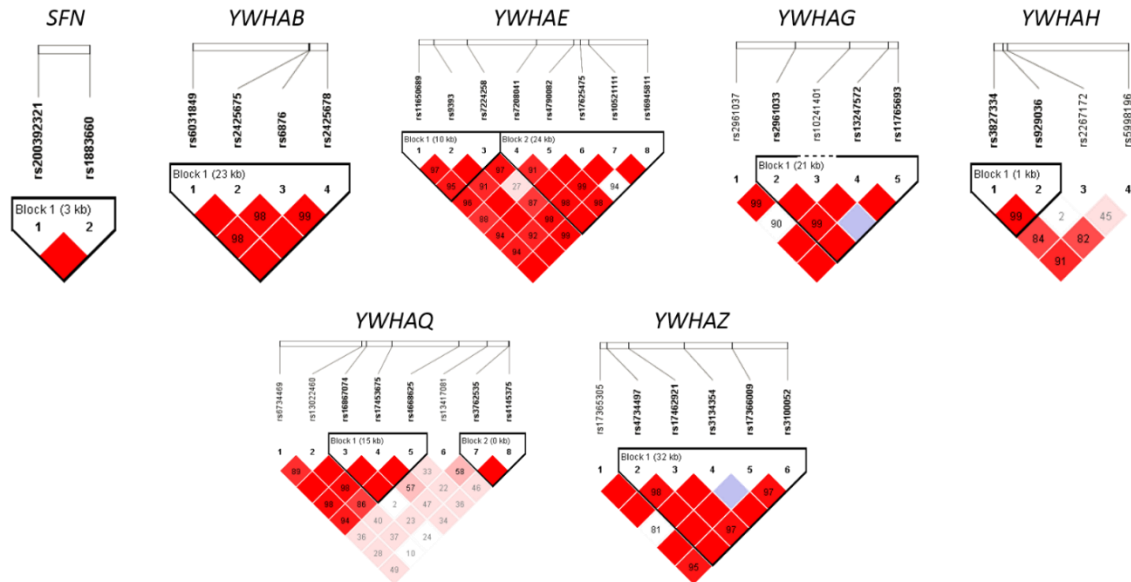


**Figure S1. Linkage disequilibrium values among the 37 tagSNPs analyzed in this study**, calculated from the whole sample (1,441 individuals with European ancestry) with the Haploview software. D' values between all the possible SNP pairs are shown.
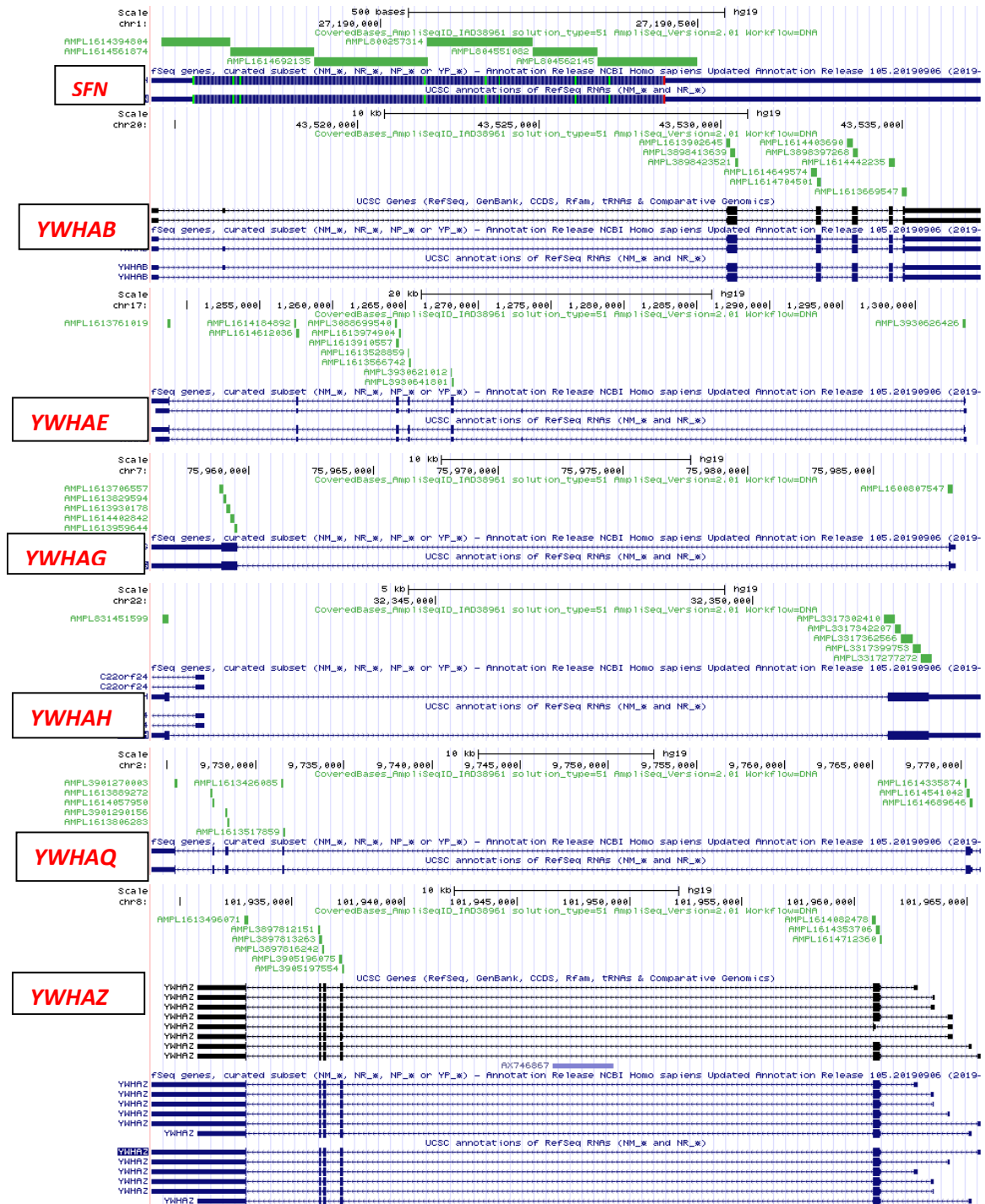
**Figure S2. The 57 amplicons used in the mutational screening are depicted in green and cover the coding regions of the 14-3-3 genes.** For each of the seven genes we show the amplicon ID name, the number of overlapping amplicons per exon and the genomic region.
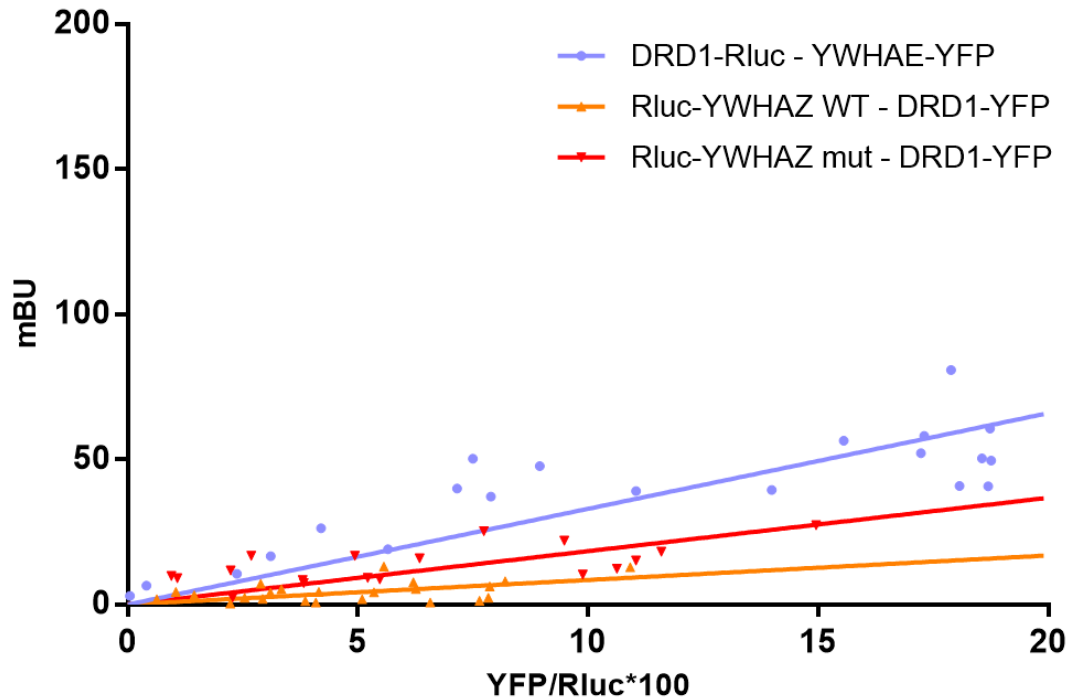
**Figure S3. Negative controls in the BRET experiments of the interaction of YWHAZ WT or YWHAZ mutant (mut) with YWHAE using D(1A) dopamine receptor as a donor (DRD1-Rluc) or an acceptor (DRD1-YFP) and adjusted to a linear regression.** mBU: BRET ratio expressed in milli-BRET units. The relative amount of BRET is given as a function of YFP/Rluc*100. YFP corresponds to the fluorescence signal due to the increasing amount of donor and Rluc corresponds to the stable luminescence signal measured at 10 minutes. BRET values shown correspond to 3-4 independent experiments.
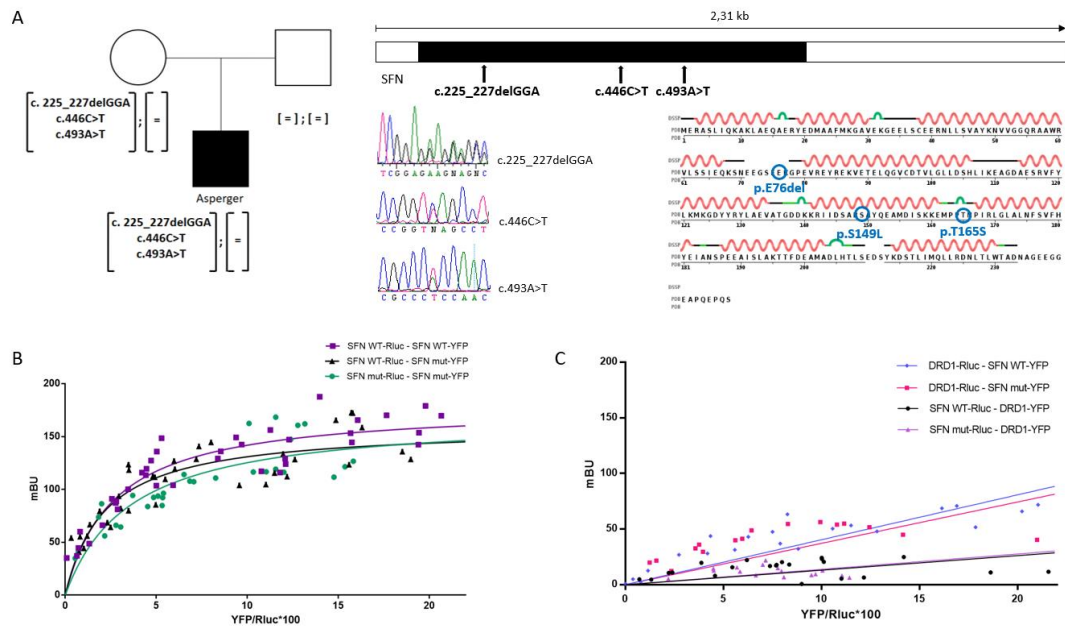
**Figure S4. Characterization of three rare inherited *SFN* variants identified in an ASD patient**
A) Pedigree of the family carrying the mutations. [=], wild-type allele and location of the mutations in the *SFN* gene, Sanger sequences corresponding to the three mutations identified in the patient, and SFN protein sequence together with the secondary structure showing the location of the amino acid changes in the mutant protein. In blue, mutations found in the patient. Protein secondary structure obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB, http://www.rcsb.org/). Protein secondary structure legend: green curve, turn; green line, bend; red, alpha helix, empty or black line, no secondary structure assigned. C) BRET assay to determine the ability to form homodimers of wild-type SFN (SFN WT) and mutant SFN (SFN mut) co-transfected in HEK293T cells. Values shown correspond to 3-7 different experiments performed. $BRET_{50}$ (SFN mut-Rluc - SFN mut-YFP: 3.57±0.89; SFN WT-Rluc - SFN mut-YFP: 2.18±0.42; SFN WT-Rluc - SFN WT-YFP: 2.66±0.43) and BRETmax (SFN mut-Rluc - SFN mut-YFP: 170.1±13.83; SFN WT-Rluc - SFN mut-YFP: 159.3±8.35; SFN WT-Rluc - SFN WT-YFP: 179.4±7.78). D) Negative controls of SFN WT and mutant (mut) interaction using D(1A) dopamine receptor as a donor (DRD1-Rluc) or an acceptor (DRD1-YFP) and adjusted to a linear regression. Values shown correspond to 2-6 different experiments performed. mBU: BRET ratio expressed in milli-BRET units. The relative amount of BRET is given as a function of YFP/Rluc*100. YFP correspond to the fluorescence signal due to the increasing amount of donor and Rluc correspond to the stable luminescence signal measured at 10 minutes.

**REFERENCES**

1. Ghorbani, S.; Fossbakk, A.; Jorge-Finnigan, A.; Flydal, M.I.; Haavik, J.; Kleppe, R. Regulation of tyrosine hydroxylase is preserved across different homo- and heterodimeric 14-3-3 proteins. *Amino Acids* **2016**, *48*, 1221–1229, doi:10.1007/s00726-015-2157-0.

2. Demontis, D.; Walters, R.K.; Martin, J.; Mattheisen, M.; Als, T.D.; Agerbo, E.; Baldursson, G.; Belliveau, R.; Bybjerg-Grauholm, J.; Bækvad-Hansen, M.; et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **2019**, *51*, 63–75, doi:10.1038/s41588-018-0269-7.

3. Tielbeek, J.J.; Johansson, A.; Polderman, T.J.C.; Rautiainen, M.R.; Jansen, P.; Taylor, M.; Tong, X.; Lu, Q.; Burt, A.S.; Tiemeier, H.; et al. Genome-wide association studies of a broad spectrum of antisocial behavior. *JAMA Psychiatry* **2017**, *74*, 1242–1250, doi:10.1001/jamapsychiatry.2017.3069.

4. Meier, S.M.; Trontti, K.; Purves, K.L.; Als, T.D.; Grove, J.; Laine, M.; Pedersen, M.G.; Bybjerg-Grauholm, J.; Bækved-Hansen, M.; Sokolowska, E.; et al. Genetic Variants Associated with Anxiety and Stress-Related Disorders: A Genome-Wide Association Study and Mouse-Model Study. *JAMA Psychiatry* **2019**, *76*, 924–932, doi:10.1001/jamapsychiatry.2019.1119.

5. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444, doi:10.1038/s41588-019-0344-8.

6. Stahl, E.A.; Breen, G.; Forstner, A.J.; McQuillin, A.; Ripke, S.; Trubetskoy, V.; Mattheisen, M.; Wang, Y.; Coleman, J.R.I.; Gaspar, H.A.; et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* **2019**, *51*, 793–803, doi:10.1038/s41588-019-0397-8.

7. Wray, N.R.; Ripke, S.; Mattheisen, M.; Trzaskowski, M.; Byrne, E.M.; Abdellaoui, A.; Adams, M.J.; Agerbo, E.; Air, T.M.; Andlauer, T.M.F.; et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **2018**, *50*, 668–681, doi:10.1038/s41588-018-0090-3.

8. Arnold, P.D.; Askland, K.D.; Barlassina, C.; Bellodi, L.; Bienvenu, O.J.; Black, D.; Bloch, M.; Brentani, H.; Burton, C.L.; Camarena, B.; et al. Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* **2018**, *23*, 1181–1188, doi:10.1038/mp.2017.154.

9. Ripke, S.; Neale, B.M.; Corvin, A.; Walters, J.T.R.; Farh, K.H.; Holmans, P.A.; Lee, P.; Bulik-Sullivan, B.; Collier, D.A.; Huang, H.; et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **2014**, *511*, 421–427, doi:10.1038/nature13595.

10. Lee, P.H.; Anttila, V.; Erneri; Won, H.; Feng, Y.-C.A.; Rosenthal, J.; Zhu, Z.; Tucker-Drob, E.M.; Nivard, M.G.; Grotzinger, A.D.; et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **2019**, *179*, 1469-1482.e11, doi:10.1016/j.cell.2019.11.020.

# Article 2. Deficiency of the *ywhaz* gene, involved in neurodevelopmental disorders, alters brain activity and behaviour in zebrafish

**Summary in Spanish: "Deficiencia del gen *ywhaz*, implicado en trastornos del neurodesarrollo, altera la actividad cerebral y el comportamiento en pez cebra"**

Variantes genéticas de riesgo en *YWHAZ*, que codifica la proteína 14-3-3ζ, han sido relacionadas con trastornos psiquiátricos, como el trastorno del espectro autista (TEA) o la esquizofrenia, y con alteraciones del neurodesarrollo en humanos y ratones. En este estudio hemos usado un modelo de pez cebra para investigar los mecanismos mediante los cuales *YWHAZ* contribuye a trastornos del neurodesarrollo. En primer lugar, observamos una expresión pan-neuronal del gen *ywhaz* durante diferentes estadíos del desarrollo, lo que sugiere un papel importante de este gen en el desarrollo neuronal. Durante la edad adulta, la expresión de *ywhaz* está restringida a las células Purkinje en el cerebelo, una región que se ha visto alterada en pacientes con TEA. A continuación, establecimos mediante la técnica de CRISPR/Cas9 una nueva línea genoanulada de pez cebra deficiente en *ywhaz*. Realizamos análisis de imagen con calcio en el cerebro completo de larvas *wild-type* y genoanuladas para *ywhaz* y encontramos diferencias en la actividad y conectividad neuronal en el romboencéfalo. Además, peces adultos genoanulados muestran niveles disminuidos de dopamina y serotonina en esta misma área cerebral y presentan alteraciones de comportamiento que se pueden revertir mediante el uso de fluoxetina y quinpirol, fármacos que modulan la neurotransmisisón serotoninérgica y dopaminérgica. En conjunto, estos resultados sugieren un rol importante del gen *ywhaz* en el establecimiento de la conectividad neuronal durante el desarrollo. Una deficiencia del gen *ywhaz* lleva a alteraciones en la neurotransmisión dopaminérgica y serotoninérgica que probablemente sean la causa de los cambios de comportamiento observados en edad adulta.

**Reference:**

Antón-Galindo E, Dalla Vecchia E, Orlandi JG, Castro G, Loza-Alvarez P, Aguado F, Norton WHJ, Cormand B, Fernàndez-Castillo N. Deficiency of the *ywhaz* gene, involved in neurodevelopmental disorders, alters brain activity and behaviour in zebrafish. To be summited to *Molecular Psychiatry*.

# Deficiency of the *ywhaz* gene, involved in neurodevelopmental disorders, alters brain activity and behaviour in zebrafish

Antón-Galindo E[1,2,3,4]\*, Dalla Vecchia E[5]\*, Orlandi JG[6], Castro G[7], Loza-Alvarez P[7], Aguado F[8], Norton WHJ[5#], Cormand B[1,2,3,4#], Fernàndez-Castillo N[1,2,3,4#]

[1]Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[2]Centro de Investigación Biomédica en Red de Enfermedades raras (CIBERER), Spain

[3]Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain

[4]Institut de recerca Sant Joan de Déu (IRSJD), Espluges de Llobregat, Spain

[5]Department of Neuroscience, Psychology and Behaviour, College of Life Sciences, University of Leicester, Leicester, UK

[6]RIKEN Center for Brain Science, Wako-shi, Saitama 351-0198, Japan

[7]The Institute of Photonic Sciences, (ICFO) Castelldefels (Barcelona), Spain

[8]Department of Cell Biology, Physiology and Immunology, University of Barcelona; Institute of Neurosciences, UB; Barcelona, Spain.

\* equally contributed to this work; # equally supervised this work

## ABSTRACT

Genetic risk variants in *YWHAZ*, encoding 14-3-3$\zeta$, have been found to contribute to psychiatric disorders such as autism spectrum disorder and schizophrenia, and have been related to an impaired neurodevelopment in humans and mice. Here, we have used a zebrafish model to further understand the mechanisms by which *YWHAZ* contribute to neurodevelopmental disorders. We first observed pan-neuronal expression of *ywhaz* during developmental stages, suggesting an important role of this gene in neural development. During adulthood *ywhaz* expression was restricted to Purkinje cells in the cerebellum, a region that shows alterations in autistic patients. We then established a novel stable *ywhaz* knockout (KO) zebrafish line using CRISPR/Cas9 genome engineering. We performed whole-brain calcium imaging in wild-type (WT) and *ywhaz* KO larvae and found altered neural activity and functional connectivity in the hindbrain. Interestingly, adult *ywhaz* KO fish also display decreased levels of dopamine and serotonin in the hindbrain and freeze when exposed to novel stimuli, a phenotype that can be reversed with fluoxetine and quinpirole, drugs that target serotonin and dopamine neurotransmission. Together, these findings suggest an important role for *ywhaz* in establishing neuronal connectivity during developmental stages. *ywhaz* deficiency leads to impaired dopamine and serotonin neurotransmission that may underlie the altered behaviour observed during adulthood.

**INTRODUCTION**

The 14-3-3 gene family codes for a highly conserved group of molecular chaperones that play important roles in biological processes and neuronal development [1]. *YWHAZ*, encoding 14-3-3ζ, is involved in neurogenesis and neuronal migration, as shown by morphological changes in the brain of 14-3-3ζ knockout mice [2–5]. These animals present behavioural alterations, that have been related to psychiatric disorders, such as schizophrenia [2, 5], including hyperactivity, impaired memory, lower anxiety and impaired sensorimotor gating.

Several studies have pointed to an association between Y*WHAZ* and psychiatric disorders. In a previous study, we found that a frameshift mutation in the *YWHAZ* gene, inherited from a mother with depression, had functional implications in two siblings diagnosed with autism spectrum disorder (ASD) and attention deficit/hyperactivity disorder (ADHD) [6]. In addition, genetic studies have associated *YWHAZ* polymorphisms to major depression and schizophrenia [6, 7]. Finally, decreased expression of the *YWHAZ* gene and 14-3-3ζ protein was reported in post-mortem brains of ASD and schizophrenia patients [6, 8, 9], and 14-3-3 protein levels are reduced in platelets and pineal glands of ASD patients [10, 11]. However, the mechanisms by which *YWHAZ* contributes to neurodevelopmental disorders remain unclear.

The zebrafish is a powerful model to study neurodevelopmental and psychiatric disorders due to several advantages. Firstly, zebrafish present a high genetic similarity to humans and are easy to manipulate genetically. Secondly, they display well-defined behaviours that can be translated to humans in some cases. Lastly, their small size and transparency during larval stages make zebrafish ideal for *in vivo* imaging studies [12, 13]. Indeed, whole-brain imaging is a recently developed technique that allows *in vivo* neuronal activity and connectivity to be investigated [14, 15]. This novel approach combined with genetic engineering stands as an excellent tool to use in specific zebrafish models to study the neural basis of human brain disorders.

In this study, we aim to investigate the role of *YWHAZ* in brain development and function using a novel mutant line. We explore the effect of loss of *ywhaz* function on neural activity and connectivity during development, and in neurotransmission and behaviour during adulthood.

**MATERIAL AND METHODS**

**Zebrafish strains, care and maintenance**

Adult zebrafish and larvae (*Danio rerio*) were maintained at 28.5°C on a 14:10 light-dark cycle following standard protocols. All experimental procedures were approved by a local Animal

Welfare and Ethical Review board (University of Leicester and Generalitat de Catalunya). AB wild-type (WT), *Tg*(*aldoca:gap43-Venus*), *Tg*(*olig2:egfp*)[vu12], *ywhaz*[-/-], albino *Tg(elavl3:GCaMP6s)* and albino *Tg(elavl3:GCaMP6s)ywhaz*[-/-] zebrafish lines were used for the experiments.

**In situ hybridization (ISH) and immunohistochemistry (IHC)**

A specific mRNA probe targeting *ywhaz* (NCBI Reference Sequence: NM_212757.2) was prepared and ISH experiments were performed in larvae and dissected adult brains of AB wild-type (WT), *Tg*(*aldoca:gap43-Venus*), *Tg*(*olig2:egfp*)[vu12] and *ywhaz*[-/-] zebrafish strains. IHC was performed in adult *Tg*(*olig2:egfp*)[vu12], *ywhaz*[-/-] and *Tg*(*aldoca*:*gap43-Venus*)[rk22] brains. Details of the procedures are described in the Supplementary methods.

**Generation *ywhaz* zebrafish knock out using CRISPR/Cas9**

The detailed CRISPR/Cas9 targeted mutagenesis protocol is described in the Supplementary methods. A synthetic guide RNA (sgRNA) targeting exon 3 of *ywhaz* was designed and cloned into the pDR274 vector (Addgene plasmid #42250) [16]. The sgRNA was transcribed using the mMESSAGE mMACHINE Kit (Life Technologies). Cas9 mRNA was transcribed *in vitro* from the pMLM3613 vector (Addgene plasmid #42251, Keith Joung) [16] using the mMESSAGE mMACHINE T7 Ultra Kit (Life Technologies). To generate the mutants, 1 nl total volume of a mixture of 25 ng/µl sgRNA and 250 ng/µl Cas9 mRNA was injected to one-cell stage embryos to form the F0 generation. Once a F0 fish carrying an interesting indel transmitted to the germline was identified, the F1 embryos born from a cross between the selected F0 founder and AB WT were raised to adulthood, screened for the mutation of interest, and subsequently in-crossed to obtain a final stable F2 homozygous mutant line.

**Gene expression analysis using Real-Time quantitative PCR (RT-qPCR)**

Total RNA was extracted from the whole brains of WT and *ywhaz*[-/-] adult zebrafish and RT-qPCR was performed on 10 brains per genotype with three replicates as described in the Supplementary methods.

**Whole-brain imaging**

*Light sheet microscopy*

Whole-brain imaging experiments were performed on 6 days-post-fertilization (dpf) albino *Tg(elavl3:GCaMP6s)* and *Tg(elavl3:GCaMP6s)ywhaz*[-/-] zebrafish larvae. Larvae were first paralyzed for 10 minutes in a 1 mg/ml $\alpha$-bungarotoxin solution (Thermofisher). They were subsequently placed inside a fluorinated ethylene-propylene (FEP) tube (0.7 mm inner diameter), with water and E3 medium, and then into a custom-made chamber to orientate the

larvae appropriately towards the objective of the light-sheet microscope (see Supplementary methods and Figure 1).
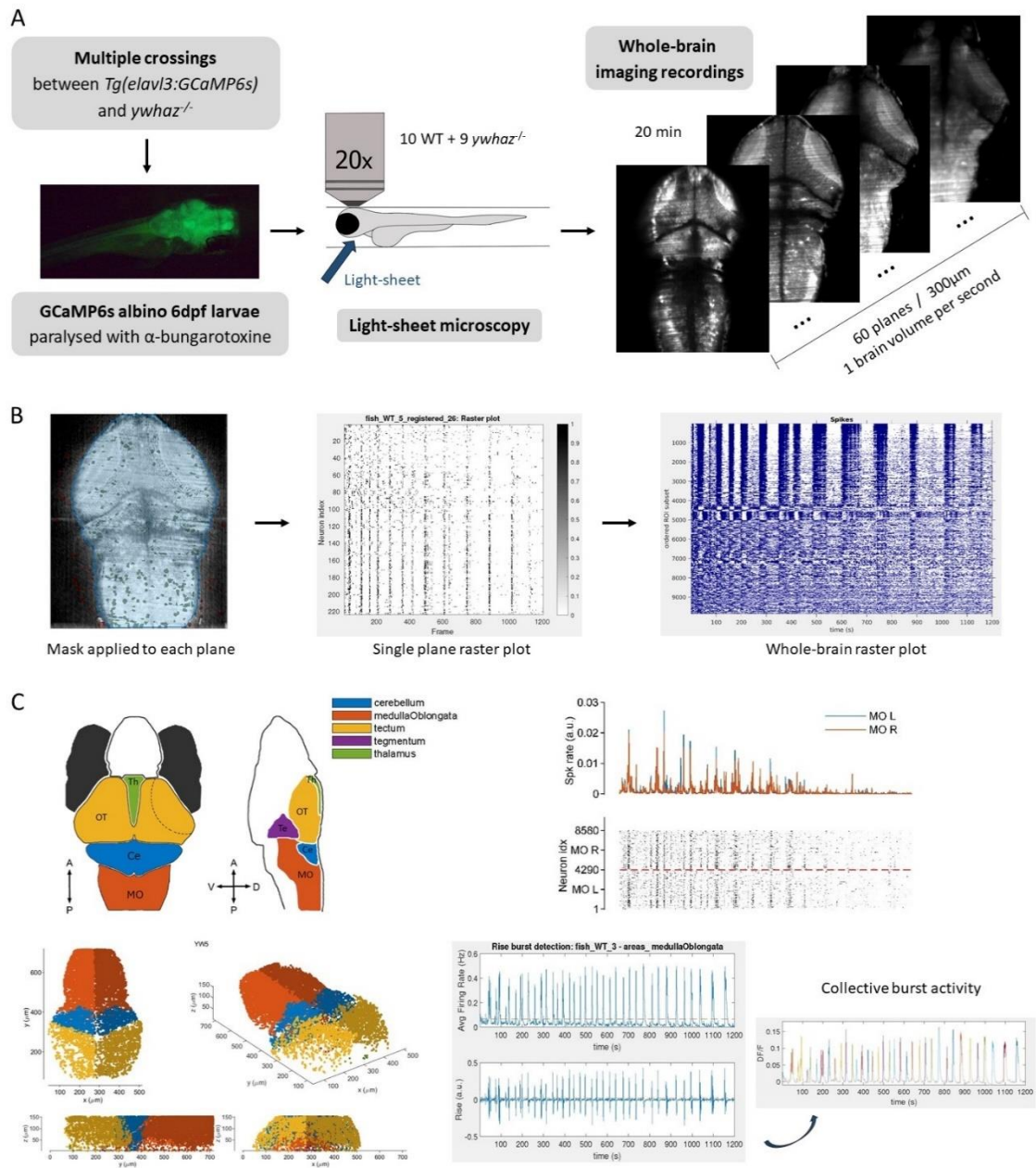


Figure 1. Whole-brain imaging methods. (A) Steps followed to perform the whole-brain imaging recordings in 6 days post-fertilization (dpf) larvae expressing GCaMP6s pan-neuronally. (B) Steps followed in the single-cell fluorescence traces extraction. First, a mask was applied to each single plane to avoid detection of fluorescence outside of the brain. Then, single cell fluorescence traces were extracted for all the neurons detected in each single plane. Finally, results from single planes were combined to obtain neuronal activity from the whole-brain of each individual during the 20min recording. ROI, region of interest. (C) On the left, brain regions explored in the analysis and plot of all the neurons detected in each brain area of one individual. In a first approach, we divided the three biggest regions (Ce, OT, and MO) by hemispheres and obtain no differences in the number of neurons or activity between hemispheres (top right). After analysing single-cell activity, we inspected collective burst activity in each of the five defined brain areas (bottom right). Ce, cerebellum; MO, medulla oblongata; OT, optic tectum; Te, tegmentum; Th, thalamus.

*Image segmentation and extraction of calcium signals*

To extract ΔF/F traces from our calcium imaging recordings, we followed the CaIman MATLAB pipeline [17]. First, volumetric data were motion corrected and separated into time series for each imaged plane. Then, individual planes were processed using constrained non-negative matrix factorization method (CNMF) [18]. We overestimated the number of ROIs, as the program discards ROIs during the refinement process, and applied a mask to avoid signal detection outside the brain. The data from each plane were then combined to produce a single dataset for an imaged larva (Figure 1).

*Activity and functional connectivity analysis*

We used Netcal (www.itsnetcal.com) to analyse fluorescence data and obtain single-cell and collective level statistics. To perform the analysis in different areas of the brain, we used MATLAB Volume Segmenter and the Z-brain atlas as a reference (http://engertlab.fas.harvard.edu/Z-Brain/) to delimitate five brain regions: thalamus, tegmentum, optic tectum, cerebellum and medulla oblongata (MO). We also performed analysis of network connectivity inside each of the delimited areas (Figure 1).

**High precision liquid chromatography (HPLC) analysis of monoamines and metabolites**

Fish were decapitated and their brains were dissected and divided into four areas: telencephalon (Tel), diencephalon (DI), optic tectum (TeO) and hindbrain (Hb). HPLC analysis for dopamine (DA), serotonin (5-HT), 3,4-dihydroxyphenylacetic acid (DOPAC), and 5-hydroxyindoleacetic acid (5-HIAA) was carried out using electrochemical detection [19] as described in the supplementary methods.

**Behavioural tests**

A battery of behavioural tests was performed in adult zebrafish (3-5 months-old) mixed groups of both sexes. Two tests to assess social behaviour were also performed on juvenile (one month-old) zebrafish. All fish were genotyped, sized-matched and maintained in groups of 15 by genotype until the day of testing. Methodological details of the tests performed are described in the Supplementary Methods.

**Drug treatments**

Fluoxetine 5 mg/L diluted in DMSO (Tocris #0927) was administered by immersion for 2 hours. Quinpirole 0.25-4 mg/L diluted in $H_2O$ (Tocris #1061) was administered by immersion for 1 hour.

**Statistical methods**

Statistical analysis of RT-qPCR, HPLC and behavioural data were performed in GraphPad Prism 6. The data sets were assessed for normality using D'Agostino-Pearson and Shaphiro-Wilk normality test. Either a Mann-Whitney or an unpaired t-test whit Welch's correction was performed to compare two groups, with corrections for multiple comparisons when needed. Statistical methods used for the brain imaging analysis are detailed in the Supplementary Methods.

**RESULTS**

**Expression of *ywhaz* during development and adulthood in WT zebrafish**

We first investigated *ywhaz* expression in the developing brain. In two to nine dpf WT larvae, *ywhaz* expression is widespread covering almost all brain areas, with a particularly strong signal in the cerebellum of whole mount embryos (Figure 2A). In contrast, in adult zebrafish, *ywhaz* expression is restricted to the cerebellum. *ywhaz* was present in the granule cell layer (GCL) of the valvula cerebelli (Va) and crista cerebellaris (CCe), most likely within the Purkinje cell layer (PCL) (Figure 2B).

We next combined *ywhaz* ISH with an anti-GFP antibody stain on adult brain sections of *Tg*(*aldoca:gap43-Venus*) [20] and *Tg*(*olig2*:*egfp*)[vu12] [21] transgenic lines, to confirm that *ywhaz* is expressed only within Purkinje cells and not in Eurydendroid cells, the zebrafish equivalent of the deep cerebellar nuclei in humans [22]. In the *Tg*(*aldoca:gap43-Venus*) line, the promoter of the *aldolase Ca* (*aldoca*) gene [20] is used as a driver and EGFP labels Purkinje cells, inhibitory neurons in the PCL. In the *Tg*(*olig2*:*egfp*)[vu12] fish cerebellum, EGFP labels Eurydendroid cells [21], excitatory neurons that are situated ventrally to Purkinje cells [23]. We found that *ywhaz* ISH staining overlaps GFP in the *Tg*(*aldoca:gap43-Venus* line), meaning that *ywhaz* is expressed within Purkinje cells and not within Eurydendroid cells at adult stages (Supplementary Figures 1 and 2).

**Generation of a *ywhaz*[-/-] zebrafish mutant line by CRISPR-Cas9 mutagenesis**

We next generated a novel *ywhaz* mutant line, as a first step towards understanding the function of this gene in neural development. We used CRISPR/Cas9 to engineer a mutation within the third exon of *ywhaz*. We selected a F0 fish carrying a 7 bp deletion (380_387delCCTGGCA) as a founder to generate a stable *ywhaz* mutant line (Supplementary Figures 3 and 4). This deletion

causes a frameshift that leads to a premature stop codon in the third exon of the unique *ywhaz* isoform. The F1 embryos born from the cross between the F0 founder and WT were raised to adulthood, genotyped, and only the F1 fish carrying the 7 bp deletion were kept (Supplementary Figure 4B).

To check whether the 7 bp deletion triggers mRNA degradation by nonsense-mediated mRNA decay (NMD) [24] we analysed *ywhaz* mRNA levels by RT-qPCR. We observed a significantly decreased level of *ywhaz* expression in *ywhaz*[-/-] compared to WT (p < 0.0001, Mann-Whitney U test, Supplementary Figure 5), which suggests that NMD degradation of the truncated *ywhaz* transcript has occurred in mutants. In addition, ISH performed in larvae and adult brains confirmed loss of *ywhaz* expression in the brain of *ywhaz*[-/-] fish, at both developmental and adult stages (Figure 2C and D).
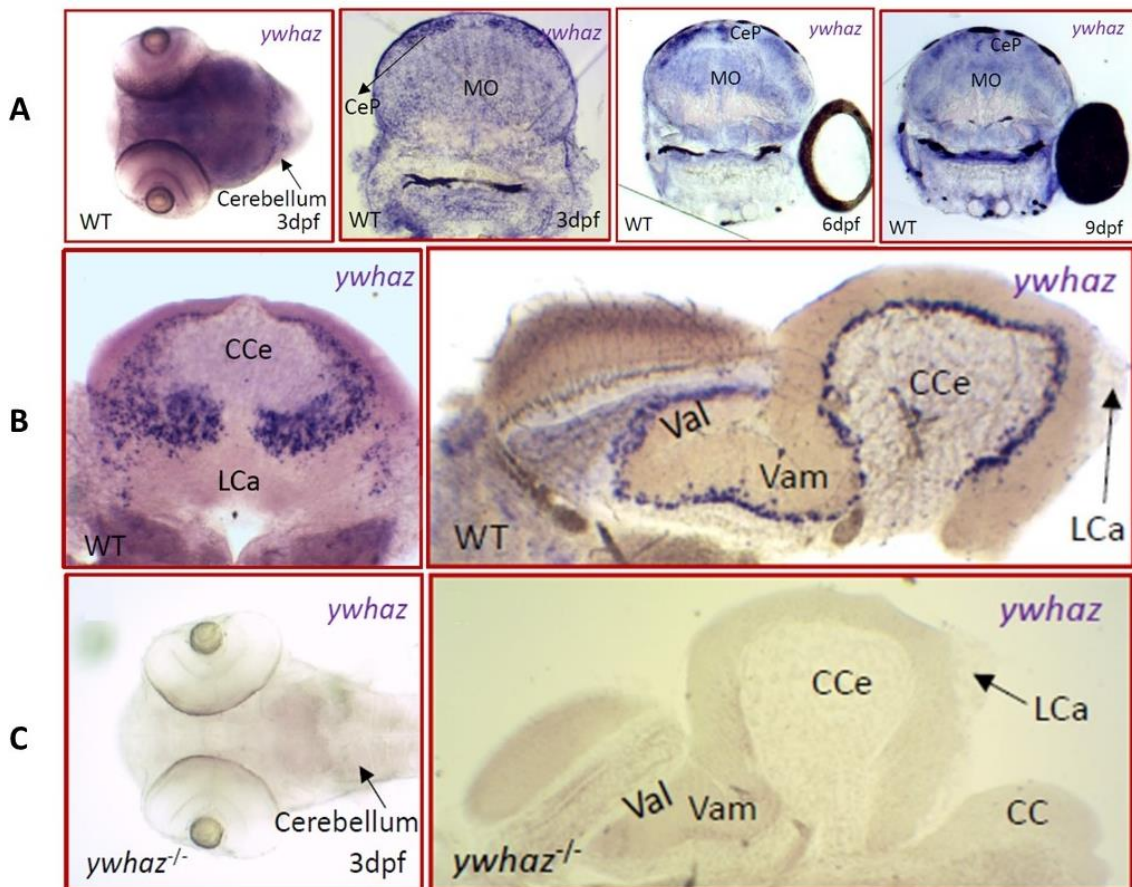


**Figure 2. Expression of *ywhaz* in zebrafish during development and adulthood in WT and KO.** (A) In WT larvae, *ywhaz* expression is widespread covering almost all brain areas, with a strongest signal in cerebellum. (B) In WT adult brains, *ywhaz* expression is restricted to the granule cell layer in the cerebellum. (C) *ywhaz* is not expressed in *ywhaz*[-/-] embryos and *ywhaz*[-/-] adult brains. CC, crista cerebellaris; CCe, corpus cerebelli; CeP, cerebellar plate; LCa, lobus caudalis cerebelli; MO, medulla oblongata; Val, lateral valvula cerebelli; Vam, medial valvula cerebelli.

**Altered spontaneous neuronal activity and functional connectivity in the hindbrain of *ywhaz<sup>-/-</sup>* larvae**

The expression pattern of *ywhaz* suggests that it may play an important role during neural development. We performed whole-brain imaging at 6 dpf to investigate changes in neural circuit function and connectivity.

When we compared WT and *ywhaz<sup>-/-</sup>* larvae we identified differences in hindbrain spontaneous activity and functional connectivity between genotypes. We found an increased number of active neurons in medulla oblongata (MO) of *ywhaz<sup>-/-</sup>* fish ($p = 0.046$, Figure 3A ), which represent a higher fraction of total active neurons ($p = 0.005$, Figure 3C). Also, *ywhaz<sup>-/-</sup>* fish cerebellar neurons represented a lower fraction of total neurons ($p = 0.018$, Figure 3C). In thalamus and tegmentum, the low number of active neurons that we could detect prevented us from properly assessing activity and connectivity in these areas (Figure 3B and D).
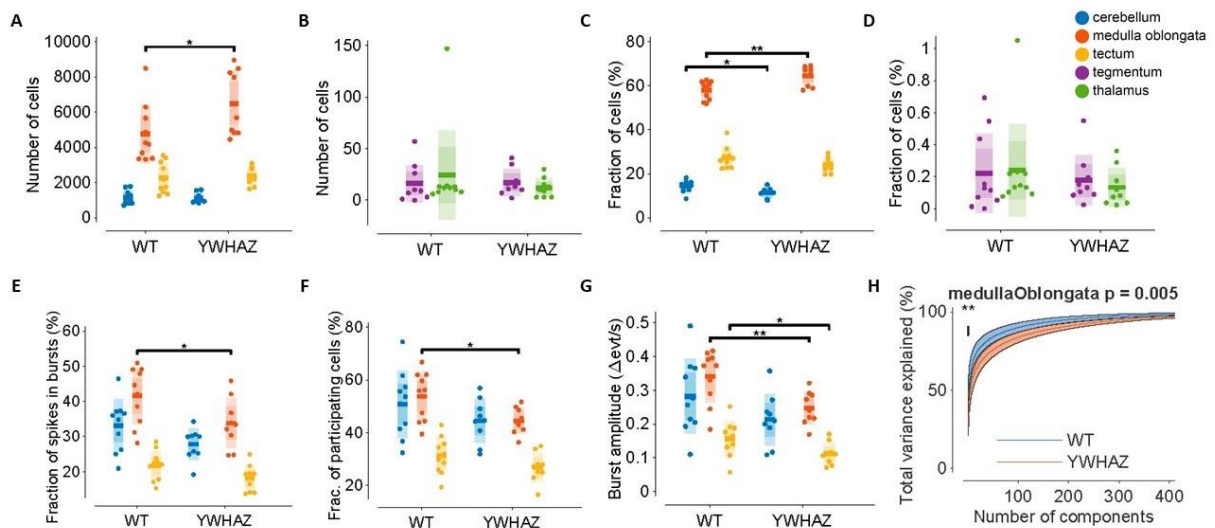


Figure 3. Alterations in neuronal activity in zebrafish *ywhaz<sup>-/-</sup>* larvae. (A and B) Number of active neurons detected in each brain area. (C and D) Fraction of total active neurons detected in each brain area. (E) Fraction of total spikes participating in single-cell bursts in each brain area. (F) Fraction of total cells participating in collective bursts in each brain area. (G) Burst amplitude in each brain area, measured as the increase of spikes (events) per second compared to basal spike activity. For A-G datasets: unpaired t-tests, each single point represents an inidvidual, the central line represents the mean, the darker bar the 95% confidence interval and the lighter bar the standard deviation. (H) Principal component analysis of the neuronal activity in the medulla oblongata (MO). The central line represents the mean and the coloured shadow represents the 95% confidence interval. * $p < 0.05$, ** $p < 0.01$. YWHAZ, *ywhaz<sup>-/-</sup>* larvae; WT, wild-type larvae. n= 10 WT and 9 *ywhaz<sup>-/-</sup>*.

Single-cell activity analysis (Supplementary Figure 6) showed that a lower fraction of total spikes participated in single-cell bursts in the MO of *ywhaz<sup>-/-</sup>* fish ($p = 0.039$, Figure 3E). In addition, collective burst analysis (Supplementary Figure 7) determined that a lower fraction of MO

neurons participate in collective bursts in *ywhaz*[-/-] fish (p = 0.016, Figure 3F), and that the collective burst amplitude in MO and tectum is lower in *ywhaz*[-/-] fish (p = 0.006 for MO and p = 0.043 for tectum, Figure 3G). Finally, a principal component analysis (PCA) demonstrated that a higher number of components is needed to explain variance in neuronal activity in MO of *ywhaz*[-/-] fish, pointing to a lower neuronal synchronization in this area in mutants (p = 0.005, Figure 3H and Supplementary Figure 8). We then explored possible alterations in connectivity inside each of the defined areas (Supplementary Figure 9). We found a higher clustering coefficient (p = 0.022, Figure 4A) and a higher global efficiency (p = 0.044, Figure 4B) in cerebellum of *ywhaz*[-/-] fish. Also, we observed a higher assortativity (p = 0.001, Figure 4C) and a higher Louvain community statistic in MO of *ywhaz*[-/-] fish (p = 0.010, Figure 4D). Finally, analyses of connectivity distribution showed that in the MO of WT fish there is a subpopulation of highly connected neurons (connected with 30-40% of the MO neurons) that is not present in the MO of *ywhaz*[-/-] fish, but no differences were found in cerebellum (Figure 4E and F).
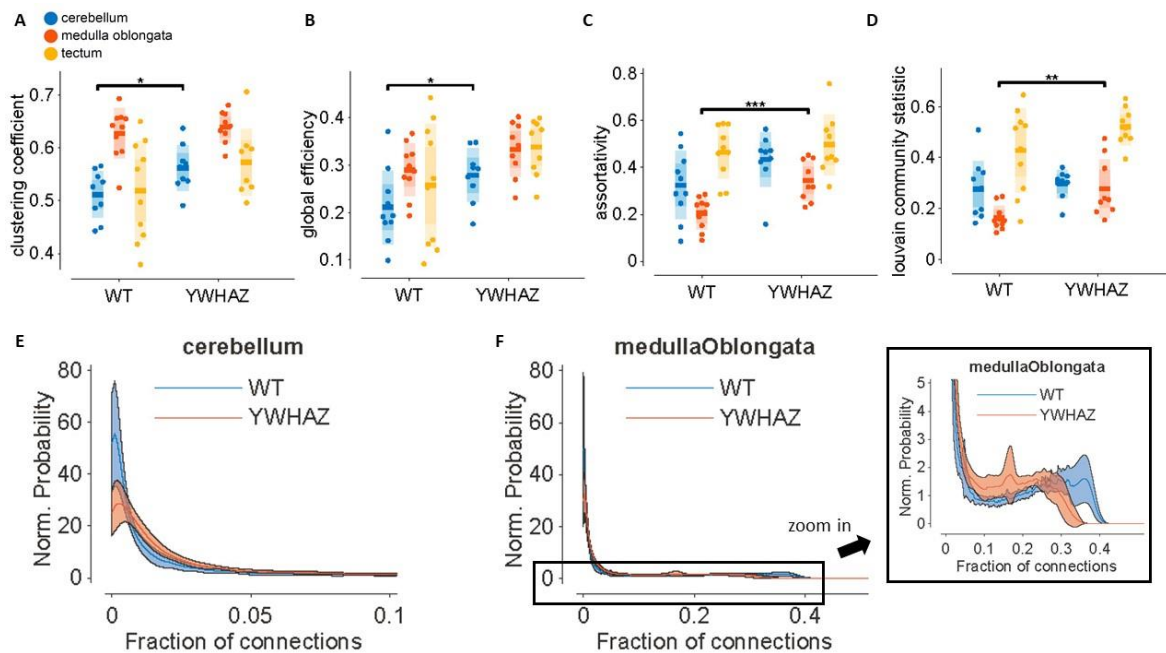


**Figure 4. Alterations in neuronal connectivity in zebrafish ywhaz[-/-] larvae.** (A) Normalized clustering coefficient in each brain area. (B) Global efficiency in each brain area. (C) Assortativity in each brain area. (D) Louvain community statistic in each brain area. For A-D datasets: unpaired t-tests, each single point represents an individual, the central line represents the mean, the darker bar the 95% confidence interval and the lighter bar the standard deviation. (E) No significant differences were found in the connectivity distribution in the cerebellum of WT and *ywhaz*[-/-] larvae. (F) Connectivity distribution in the medulla oblongata (MO) of WT and *ywhaz*[-/-] larvae. A subpopulation of highly connected neurons exists in the MO of WT larvae that is absent in *ywhaz*[-/-] larvae. For E and F: the central line represents the mean and the coloured shadow represents the 95% confidence interval. YWHAZ, *ywhaz*[-/-] larvae; WT, wild-type larvae. n= 10 WT and 9 *ywhaz*[-/-]. * p < 0.05, ** p < 0.01, *** p < 0.001.

All together, these results point to a higher clustering and more effective connectivity in the cerebellum of $ywhaz^{-/-}$ fish, and to the presence of more isolated neuronal communities in the MO of $ywhaz^{-/-}$ fish, which generates a lower collective burst activity and synchrony in this area. Spontaneous synchronized neuronal communication plays an important role during early development in establishing the mature brain circuitry, not only in humans but also in other vertebrates including zebrafish [25–28]. We hypothesize that the alterations in this spontaneous activity may affect neuronal migration and wiring and have a long-term impact in neurotransmission.

**Analysis of monoamines and their metabolites levels**

We next investigated whether the observed alterations in neural activity and connectivity caused by a loss of function of $ywhaz$ gene during development might affect neurotransmitter signalling in adults. We performed HPLC to measure the basal levels of several neurotransmitters: DOPAC, DA, 5HIAA and 5-HT, in the brain of $ywhaz^{-/-}$ and WT adult fish. We found a significant reduction of DA and 5-HT levels in the hindbrain of $ywhaz^{-/-}$ (p = 0.0063 and p = 0.0026 respectively, Figure 5A). No further alterations were found in other areas of the brain (Supplementary Figure 10A-C) nor in the breakdown of 5-HT to 5-hydroxyindoleacetic acid (5HIAA), or dopamine to homovanillic acid (HVA) and dihydroxyphenolacetic acid (DOPAC) (Supplementary Figure 10D-E).

**Effects of loss of $ywhaz$ function on gene expression**

The reduced levels of DA and 5-HT in mutants suggests that $ywhaz$ may influence the synthesis of these neurotransmitters. Tyrosine hydroxylase (TH) and Tryptophan hydroxylase (TPH), rate-limiting enzymes in the biosynthesis of DA, are known to be regulated by 14-3-3 proteins [29]. We therefore first measured the expression of genes coding for the DA and 5-HT synthesis enzymes TH and TPH in adult fish. There was an increase in *tryptophan hydroxylase 2* (*tph2*) and tyrosyne hydroxylase 1 (*th1*) expression in $ywhaz^{-/-}$ compared to WT, although differences in *th1* did not overcome multiple test corrections. *tph1a*, *tph1b* and *th2* showed all a low expression in both genotypes (Figure 5A). We further examined the transcription of genes coding for proteins involved in the dopaminergic neurotransmitter pathway: the *dopamine transporter solute carrier family 6 member 3* (*slc6a3*), *dopamine receptor 1* (*drd1*), *dopamine receptor 2a* (*drd2a*), *dopamine receptor 2b* (*drd2b*) and *dopamine receptor 4* (*drd4*). The expression of the two isoforms *drd2a and drd2b* was significantly increased in $ywhaz^{-/-}$ compared to WT (Figure 5A).
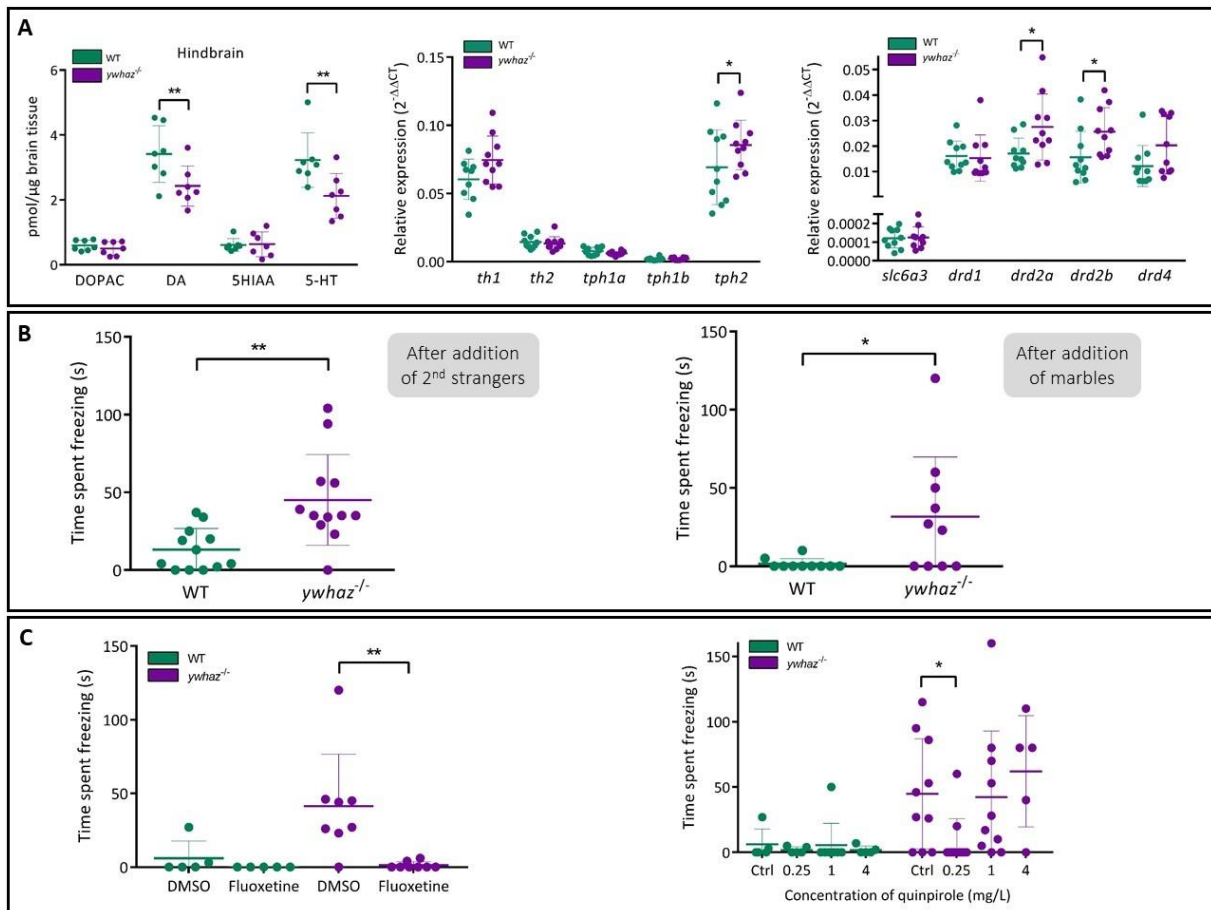
**Figure 5. Alterations in the monoamine neurotransmission in the hindbrain of adult KO.** (A) On the left, high precision liquid chromatography in the hindbrain of WT and *ywhaz*[-/-] adult zebrafish. DA and 5-HT levels are decreased in the hindbrain of *ywhaz*[-/-] compared to WT. DA, dopamine; DOPAC, 3,4-dihydroxyphenylacetic acid; 5-HIAA, 5-hydroxyindoleacetic acid; 5-HT, 5-hydroxytryptamine. n = 7 WT, n = 7 ywhaz[-/-]. In the middle, relative expression profile of *tyrosine hydroxylase 1* (*th1*), *tyrosine hydroxylase 2* (*th2*), *tryptophan hydroxylase 1a* (*tph1a*), *tryptophan hydroxylase 1b* (*tph1b*) and *tryptophan hydroxylase 2* (*tph2*), normalised to the reference gene ribosomal protein L13 (*rpl13*). *ywhaz*[-/-] fish have an increased level of *tph2* expression. n = 10 WT, n = 10 *ywhaz*[-/-]. On the right, relative expression profile *of solute carrier family 6 (neurotransmitter transporter), member 3* (*slc6a3*), *dopamine receptor 1* (*drd1*), *dopamine receptor 2a* (*drd2a*), *dopamine receptor 2b* (*drd2b*) and *dopamine receptor 4* (*drd4*), normalised to the reference gene *elongation factor 1a* (*elf1a*). *ywhaz*[-/-] have an increased level of *drd2a* and *drd2b* expression. n = 10 WT, n = 10 *ywhaz*[-/-]. Multiple t-tests with Holm-Sidak correction for multiple comparisons. (B) Time spent freezing during the first two minutes after the addition of the second group of strangers or marbles in the behavioural setup during the second step of the visually mediated social preference test. Strangers: Unpaired t-test with Welch's correction; n = 12. Marbles: Mann-Whitney U test; n = 10. (C) Treatment with 5 mg/L fluoxetine and 0.25 mg/L quinpirole rescues the freezing phenotype in *ywhaz*[-/-]. On the left, time spent freezing after the addition of the second group of unfamiliar fish in WT or *ywhaz*[-/-] fish treated with 5 mg/L fluoxetine or DMSO. Multiple t-tests with Holm-Sidak correction for multiple comparisons. n = 5 WT, n = 8 *ywhaz*[-/-]. On the right, time spent freezing after the addition of the second group of unfamiliar fish in WT or *ywhaz*[-/-] fish treated with different concentrations of quinpirole. Kruskal-Wallis test with Dunn's multiple comparisons. n = 5 WT and n = 10 *ywhaz*[-/-] for Ctrl, 0.25 mg/L and 1 mg/L; and n = 5 *ywhaz*[-/-] for 4 mg/L. ** p < 0.01, * p < 0.05. Mean ± standard deviation.

**Pharmacological reversion of *ywhaz*⁻/⁻ altered behavioural phenotype**

A battery of behavioural tests was performed in adult WT and *ywhaz*⁻/⁻ mutants to characterize possible alterations in behaviour due to *ywhaz* deficiency (Supplementary Figure 11). We only found differences between genotypes in the visually-mediated social preference (VMSP) test, used to assess social behaviour. During the social preference step, both genotypes spent most of the time swimming close to the first group of strangers ($p < 0.0001$ for both WT and *ywhaz*⁻/⁻ ; Supplementary Figure 11D). However, during the preference for social novelty step, WT switched preference to the second group of unfamiliar fish ($p = 0.0475$, Supplementary Figure 11D) whereas *ywhaz*⁻/⁻ fish did not show preference between the two groups of strangers ($p = 0.98$, Supplementary Figure 11D). Instead, we found that immediately after the addition of the second group of unfamiliar fish, *ywhaz*⁻/⁻ mutants froze significantly more than WT ($p = 0.0035$; Figure 5B). To test whether it was the live stimulus that caused this phenotype, we repeated the test by adding marbles instead of the second group of unfamiliar fish and observed the same freezing behaviour in *ywhaz*⁻/⁻ mutants ($p = 0.0251$, Figure 5B). Contrary to adults, juvenile ywhaz⁻/⁻ fish did not show an altered social behaviour: they did not show any increase in freezing behaviour in the VMSP test ($p > 0.99$, Supplementary Figure 12A-C), nor an altered cluster score in the shoaling test for social interaction in a group (Supplementary Figure 12D).

HPLC and qPCR results suggested that alterations to 5-HT and DA signalling may underlie the behavioural phenotype of ywhaz⁻/⁻. We therefore used fluoxetine, a selective serotonin reuptake inhibitor (SSRI) [30], and quinpirole, a selective D2-like receptor agonist [31] to investigate the connection between 5-HT, DA, and the behavioural phenotype observed in *ywhaz*⁻/⁻ in the VMSP test. Treatment with 5 mg/L fluoxetine and treatment with 0.25 mg/L quinpirole significantly decreased the time *ywhaz*⁻/⁻ spent freezing after the addition of the second group of unfamiliar fish ($p = 0.006$ and $p= 0.0468$, respectively, Figure 5C) without affecting WT behaviour. However, a higher concentration of quinpirole, 1 or 4 mg/L, did not reverse the freezing behaviour ($p > 0.99$ for both concentrations, Figure 5C).

Several behavioural tests were repeated in a second group of adult fish and an increased freezing behaviour was observed in ywhaz⁻/⁻ fish in all the tests performed (Supplementary Figure 13). Additionally, a third batch of behavioural tests was performed in a different setup, but strong differences in the WT behaviour of this batch, probably due to environmental effects, prevent us from comparing the results obtained in this batch with the previous batches (Supplementary Figure 14).

**DISCUSSION**

Functional and genetic studies had previously suggested a role of *YWHAZ* gene in neurodevelopmental disorders such as ASD or schizophrenia [2, 5, 6, 32]. In this study, we have used zebrafish to investigate *ywhaz* function in neural development and adult behaviour. We found that *ywhaz* deficiency resulted in an altered hindbrain connectivity during larval stages and a neurotransmitter imbalance during adulthood leading to alterations in behaviour.

We observed a pan-neuronal expression of *ywhaz* in larval stages, which suggests that *ywhaz* may be involved in a wide range of functions during zebrafish neurodevelopment. Indeed, the important function of *Ywhaz* in neurogenesis and neuronal differentiation was previously demonstrated in mice KO models [1–5], and disrupted neurogenesis is a known risk factor for neurodevelopmental disorders [33]. Although expressed pan-neuronally, *ywhaz* presents a stronger expression in larvae cerebellum, a brain region whose dysfunction has been related to neurodevelopmental disorders [34, 35]. Interestingly, a decreased *YWHAZ* expression was reported in the cerebellum of ASD patients [6]. In adults, *ywhaz* is specifically expressed in Purkinje cells (PC) in the cerebellum. Post-mortem studies have shown a reduction in PC density in the brain of autistic patients, and *Tsc1* mutant mice with a decrease in PC functioning show autistic-like behaviours [36, 37]. *ywhaz* is therefore expressed in brain regions that play a crucial role in neurodevelopment and whose dysfunction is related to neurodevelopmental disorders.

Early spontaneous synchronized burst activity is essential for the correct assembly of neural circuits during development [25, 38]. Indeed, neuronal activity controls apoptosis and future connectivity in the developing cortex of mice [39], and a recent article reviewed the relevance of spontaneous activity in the larval zebrafish optic tectum and its implication in the response to sensory information [26]. In addition, magnetic resonance studies have described an altered functional connectivity in patients diagnosed with ASD or schizophrenia [40–44]. Here, *ywhaz*[-/-] larvae present a decreased burst activity and synchronization in the hindbrain, the region where *ywhaz* showed a higher expression in WT animals. We then identified decreased levels of DA and 5-HT in the hindbrain of *ywhaz*[-/-] adults, suggesting long-term alterations in brain function. These findings suggest that *ywhaz* is involved in establishing brain connectivity during development and that this impaired connectivity may contribute to the subsequent neurotransmitter alterations found in adults.

Among all 14-3-3 isoforms, *YWHAZ* plays the most important role in DA synthesis [45]. Indeed, *YWHAZ* is known to regulate the function of TH and TPH, rate-limiting enzymes in the biosynthesis of DA [29]. In a previous work, we demonstrated that a disrupting mutation in

*YWHAZ* produced the loss of affinity of the protein for TH [6]. Therefore, *ywhaz* deficiency may contribute to a decrease in TH activity and a subsequent reduction in DA levels. These results are in line with the alterations in DA system previously reported in patients with neurodevelopmental disorders such as ASD, ADHD or schizophrenia [46–48]. The increased *drd2a* and *drd2b* levels we reported in *ywhaz* mutants may be due to compensatory mechanisms to overcome DA depletion. Interestingly, schizophrenia patients have a higher density of DRD2 receptors and all antipsychotic drugs used today are DRD2 antagonists [49, 50]. In addition, upregulation of *tph2* may be a mechanism to compensate for the depletion of 5-HT we found in the hindbrain of *ywhaz$^{-/-}$* adults. In line with these results, altered levels of 5-HT have been reported in ASD and schizophrenia patients [51, 52].

Finally, treatment with fluoxetine, a serotonin reuptake inhibitor, and quinpirole, a selective DRD2-like receptor agonist, were able to rescue the abnormal neophobic freezing behaviour observed in *ywhaz* mutants. Similarly, in a previous study, behavioural alterations were rescued in *Ywhaz$^{-/-}$* mice using the antipsychotic drug clozapine, an antagonist of DA and 5-HT receptors used as a medication for schizophrenia [53]. 5-HT and DA are involved in sensory processing and social cognition [54, 55], and DA plays an important role in social reward [56]. Given 5-HT and DA function in behaviour, we hypothesize that 5-HT and DA alterations in mutant fish are responsible for the exaggerated response to novel stimuli present in *ywhaz$^{-/-}$* adults.

Several strengths and limitations of this study should be discussed. First, whole-brain imaging experiments were performed in 6 dpf larvae whereas neurotransmitter levels and behaviour were investigated in adult fish, due to the impossibility of performing all these experiments at the same age. Even though this constitutes a limitation, it brings useful complementary information to analyse the effect of *ywhaz* deficiency at different levels and ages. In addition, our whole-brain imaging setup did not permit us to analyse neuronal activity in ventral brain regions. Therefore, it would be interesting to perform similar imaging recordings in an adapted setup that allows to scan deep brain areas and to analyse brain activity in response to stimuli. Finally, the differences observed in the behavioural phenotype between batches of experiments may be due to environmental differences (between setups) and to (epi)genetic changes between fish generations. Indeed, a previous study reported differences between *Ywhaz$^{-/-}$* mice with a different genetic background [2], and it is well described that environment plays an important role in the onset or phenotypical expression of psychiatric and neurodevelopmental disorders.

In conclusion, our findings highlight the important role of *YWHAZ* in neurodevelopment and shed light on the neurobiological mechanisms underlying its contribution to

neurodevelopmental disorders. Pharmacological recovery of altered behaviour of *ywhaz*[-/-] fish provide some clues for the use of specific treatments to revert the associated symptomatology to neuropsychiatric disorders, such as ASD or schizophrenia.

## REFERENCES

1. Cornell B, Toyo-oka K. 14-3-3 proteins in brain development: Neurogenesis, neuronal migration and neuromorphogenesis. Front Mol Neurosci. 2017;10:318.

2. Xu X, Jaehne EJ, Greenberg Z, McCarthy P, Saleh E, Parish CL, et al. 14-3-3ζ deficient mice in the BALB/c background display behavioural and anatomical defects associated with neurodevelopmental disorders. Sci Rep. 2015;5:12434.

3. Toyo-Oka K, Wachi T, Hunt RF, Baraban SC, Taya S, Ramshaw H, et al. 14-3-3E and Z Regulate Neurogenesis and Differentiation of Neuronal Progenitor Cells in the Developing Brain. J Neurosci. 2014;34:12168–12181.

4. Jaehne EJ, Ramshaw H, Xu X, Saleh E, Clark SR, Schubert KO, et al. In-vivo administration of clozapine affects behaviour but does not reverse dendritic spine deficits in the 14-3-3ζ KO mouse model of schizophrenia-like disorders. Pharmacol Biochem Behav. 2015;138:1–8.

5. Cheah PS, Ramshaw HS, Thomas PQ, Toyo-Oka K, Xu X, Martin S, et al. Neurodevelopmental and neuropsychiatric behaviour defects arise from 14-3-3ζ deficiency. Mol Psychiatry. 2012;17:451–466.

6. Torrico B, Antón-Galindo E, Fernàndez-Castillo N, Rojo-Francàs E, Ghorbani S, Pineda-Cirera L, et al. Involvement of the 14-3-3 Gene Family in Autism Spectrum Disorder and Schizophrenia: Genetics, Transcriptomics and Functional Analyses. J Clin Med. 2020;9:1851.

7. Jia Y, Yu X, Zhang B, Yuan Y, Xu Q, Shen Y, et al. An association study between polymorphisms in three genes of 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein) family and paranoid schizophrenia in northern Chinese population. Eur Psychiatry. 2004;19:377–379.

8. Middleton FA, Peng L, Lewis DA, Levitt P, Mirnics K. Altered expression of 14-3-3 in the prefrontal cortex of subjects with schizophrenia. Neuropsychopharmacology. 2005;30:974–983.

9. English JA, Pennington K, Dunn MJ, Cotter DR. The neuroproteomics of schizophrenia. Biol Psychiatry. 2011;69:163–172.

10. Pagan C, Delorme R, Callebert J, Goubran-Botros H, Amsellem F, Drouot X, et al. The serotonin-N-acetylserotonin-melatonin pathway as a biomarker for autism spectrum disorders. Transl Psychiatry. 2014;4:e479.

11. Pagan C, Goubran-Botros H, Delorme R, Benabou M, Lemière N, Murray K, et al. Disruption of melatonin synthesis is associated with impaired 14-3-3 and miR-451 levels in patients with autism spectrum disorders. Sci Rep. 2017;7:2096.

12. Kalueff A V., Stewart AM, Gerlai R. Zebrafish as an emerging model for studying complex brain disorders. Trends Pharmacol Sci. 2014;35:63–75.

13. Norton W. Towards developmental models of psychiatric disorders in zebrafish. Front Neural Circuits. 2013;7:1–12.

14. Ahrens MB, Orger MB, Robson DN, Li JM, Keller PJ. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat Methods. 2013;10:413–420.

15. Vanwalleghem GC, Ahrens MB, Scott EK. Integrative whole-brain neuroscience in larval zebrafish. Curr Opin Neurobiol. 2018;50:136–145.

16. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol. 2013;31:227–229.

17. Giovannucci A, Friedrich J, Gunn P, Kalfon J, Brown BL, Koay SA, et al. CaImAn an open source tool for scalable calcium imaging data analysis. Elife. 2019;8:1–45.

18. Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, et al. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron. 2016;89:285.

19. Young AMJ. Increased extracellular dopamine in nucleus accumbens in response to unconditioned and

conditioned aversive stimuli: Studies using 1 min microdialysis in rats. J Neurosci Methods. 2004;138:57–63.

20. Ahn AH, Dziennis S, Hawkes R, Herrup K. The cloning of zebrin II reveals its identity with aldolase C. Development. 1994;120:2081–2090.

21. McFarland KA, Topczewska JM, Weidinger G, Dorsky RI, Appel B. Hh and Wnt signaling regulate formation of olig2+ neurons in the zebrafish cerebellum. Dev Biol. 2008;318:162–171.

22. Hibi M, Shimizu T. Development of the cerebellum and cerebellar neural circuits. Dev Neurobiol. 2012;72:282–301.

23. Biechl D, Dorigo A, Köster RW, Grothe B, Wullimann MF. Eppur Si muove: Evidence for an external granular layer and possibly transit amplification in the teleostean cerebellum. Front Neuroanat. 2016;10.

24. Karousis ED, Nasif S, Mühlemann O. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. Wiley Interdiscip Rev RNA. 2016;7:661–682.

25. Molnár Z, Luhmann HJ, Kanold PO. Transient cortical circuits match spontaneous and sensory-driven activity during development. Science (80- ). 2020;370.

26. Marachlian E, Avitan L, Goodhill GJ, Sumbre G. Principles of functional circuit connectivity: Insights from spontaneous activity in the zebrafish optic tectum. Front Neural Circuits. 2018;12:1–8.

27. Avitan L, Pujic Z, Mölter J, Van De Poll M, Sun B, Teng H, et al. Spontaneous Activity in the Zebrafish Tectum Reorganizes over Development and Is Influenced by Visual Experience. Curr Biol. 2017;27:2407-2419.e4.

28. Momose-Sato Y, Sato K. Development of spontaneous activity in the avian hindbrain. Front Neural Circuits. 2016;10:1–8.

29. Aitken A. 14-3-3 proteins: A historic overview. Semin Cancer Biol. 2006;16:162–172.

30. Vaswani M, Linda FK, Ramesh S. Role of selective serotonin reuptake inhibitors in psychiatric disorders: A comprehensive review. Prog Neuro-Psychopharmacology Biol Psychiatry. 2003;27:85–102.

31. Millan MJ, Maiofiss L, Cussac D, Audinot V, Boutin JA, Newman-Tancredi A. Differential actions of antiparkinson agents at multiple classes of monoaminergic receptor. I. A multivariate analysis of the binding profiles of 14 drugs at 21 native and cloned human receptor subtypes. J Pharmacol Exp Ther. 2002;303:791–804.

32. Toma C, Torrico B, Hervás A, Valdés-Mas R, Tristán-Noguero A, Padillo V, et al. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. Mol Psychiatry. 2014;19:784–790.

33. Packer A. Neocortical neurogenesis and the etiology of autism spectrum disorder. Neurosci Biobehav Rev. 2016;64:185–195.

34. Stoodley CJ. The Cerebellum and Neurodevelopmental Disorders. Cerebellum. 2016;15:34–37.

35. van der Heijden ME, Gill JS, Sillitoe R V. Abnormal Cerebellar Development in Autism Spectrum Disorders. Dev Neurosci. 2021:1–10.

36. Tsai PT, Hull C, Chu Y, Greene-Colozzi E, Sadowski AR, Leech JM, et al. Autistic-like behaviour and cerebellar dysfunction in Purkinje cell Tsc1 mutant mice. Nature. 2012;488:647–651.

37. Skefos J, Cummings C, Enzer K, Holiday J, Weed K, Levy E, et al. Regional Alterations in Purkinje Cell Density in Patients with Autism. PLoS One. 2014;9:e81255.

38. Kirkby LA, Sack GS, Firl A, Feller MB. A role for correlated spontaneous activity in the assembly of neural circuits. Neuron. 2013;80:1129–1144.

39. Blanquie O, Yang JW, Kilb W, Sharopov S, Sinning A, Luhmann HJ. Electrical activity controls area-specific expression of neuronal apoptosis in the mouse developing cerebral cortex. Elife. 2017;6:1–21.

40. O'Reilly C, Lewis JD, Elsabbagh M. Is functional brain connectivity atypical in autism? A systematic review of EEG and MEG studies. PLoS One. 2017;12:1–28.

41. Rane P, Cochran D, Hodge SM, Haselgrove C, Kennedy DN, Frazier JA. Connectivity in Autism: A Review of MRI Connectivity Studies. Harv Rev Psychiatry. 2015;23:223–244.

42. Li X, Zhang K, He X, Zhou J, Jin C. Structural , Functional , and Molecular Imaging of Autism Spectrum Disorder. Neurosci Bull. 2021. 2021. https://doi.org/10.1007/s12264-021-00673-0.

43. Sheffield JM, Barch DM. Cognition and resting-state functional connectivity in schizophrenia. Neurosci Biobehav Rev. 2016;61:108–120.

44. Van Den Heuvel MP, Fornito A. Brain networks in schizophrenia. Neuropsychol Rev. 2014;24:32–48.

45. Wang J, Lou H, Pedersen CJ, Smith AD, Perez RG. 14-3-3ζ contributes to tyrosine hydroxylase activity in MN9D cells: Localization of dopamine regulatory proteins to mitochondria. J Biol Chem. 2009;284:14011–14019.

46. Kesby JP, Eyles DW, McGrath JJ, Scott JG. Dopamine, psychosis and schizophrenia: The widening gap between basic and clinical neuroscience. Transl Psychiatry. 2018;8:30.

47. Pavăl D. A Dopamine Hypothesis of Autism Spectrum Disorder. Dev Neurosci. 2017;39:355–360.

48. Tripp G, Wickens JR. Neurobiology of ADHD. Neuropharmacology. 2009;57:579–589.

49. Zakzanis KK, Hansen KT. Dopamine D2 densities and the schizophrenic brain. Schizophr Res. 1998;32:201–206.

50. Kestler LP, Walker E, Vega EM. Dopamine receptors in the brains of schizophrenia patients: A meta-analysis of the findings. Behav Pharmacol. 2001;12:355–371.

51. Gabriele S, Sacco R, Persico AM. Blood serotonin levels in autism spectrum disorder: A systematic review and meta-analysis. Eur Neuropsychopharmacol. 2014;24:919–929.

52. Muck-Seler D, Pivac N, Mustapic M, Crncevic Z, Jakovljevic M, Sagud M. Platelet serotonin and plasma prolactin and cortisol in healthy, depressed and schizophrenic women. Psychiatry Res. 2004;127:217–226.

53. Ramshaw H, Xu X, Jaehne EJ, McCarthy P, Greenberg Z, Saleh E, et al. Locomotor hyperactivity in 14-3-3ζ KO mice is associated with dopamine transporter dysfunction. Transl Psychiatry. 2013;3:e327.

54. Jacob SN, Nienborg H. Monoaminergic Neuromodulation of Sensory Processing. Front Neural Circuits. 2018;12:1–17.

55. Fernández M, Mollinedo-Gajate I, Peñagarikano O. Neural Circuits for Social Cognition: Implications for Autism. Neuroscience. 2018;370:148–162.

56. Rademacher L, Schulte-Rüther M, Hanewald B, Lammertz S. Reward: From basic reinforcers to anticipation of social cues. Curr. Top. Behav. Neurosci., vol. 30, Springer Verlag; 2016. p. 207–221.

## SUPPLEMENTARY MATERIAL

## 1.  SUPPLEMENTARY METHODS

**Generation of *ywhaz* zebrafish knock out using CRISPR/Cas9**

*Design of the target sgRNA*

A synthetic guide RNA (sgRNA) was designed using the online available software CHOPCHOP [1,2]. We selected a target sequence containing the PAM motif within the third exon of *ywhaz* and we designed the sgRNA based on this sequence: 5'- GGGTGACTATTACCGCTACC -3' (Supplementary Figure 3).

We extracted genomic DNA from 10 adult zebrafish (5 males and 5 females) and amplified the area around the CRISPR target site by PCR (forward (Fw) primer 5'- TGACCTGGTTTCTGAGCTGA -3' and reverse (Rv) primer 5'- TGCTGAACATCAAAGACCATCT -3'). Then, we used the T7 endonuclease I (T7EI) assay to check if any mutation was present in and around the target site. 5 µl of PCR product derived from each male was mixed with 5 µl of genomic DNA from each female and the samples were diluted to a final volume of 20 µl. We denatured the samples and then reanneal them with the addition of 2 µl of NEBuffer 2 (New England Biolabs) using a thermocycler and the following protocol (95°C, 5 min; 95-85°C at −2°C/s; 85-25°C at −0.1°C/s; hold at 4°C). Hybridized PCR products were then treated with 2 U T7EI at 37°C for 1 hour in a reaction volume of 20 µl. The action of the T7EI was visualized by gel electrophoresis and the absence of SNPs in this amplified fragment was confirmed.

*Production of sgRNA*

We used the pDR274 vector, a gift from Keith Joung (Addgene plasmid #42250) [3], to generate the template for sgRNA transcription. In order to construct a plasmid containing the designed sgRNA, the vector was digested with *BsaI* restriction enzyme (New England Biolabs), dephosphorylated using Antarctic phosphatase (New England Biolabs), and purified from a gel. Two oligonucleotides 5'- TAGGGTGACTATTACCGCTACC -3' and 5'- AAACGGTAGCGGTAATAGTCAC-3' were designed to generate sgRNA. Oligonucleotides were

phosphorylated using T4 Polynucleotide Kinase (New England Biolabs) and annealed by incubation at 95°C for 5 minutes followed by slowly decreasing the temperature. The annealed oligonucleotides were then cloned into the vector backbone using the *BsaI* restriction site.

The sgRNA was transcribed using the mMESSAGE mMACHINER Kit (Life Technologies) and the *DraI*-digested gRNA expression vector as a template. The sgRNA was DNase treated and precipitated with ammonium acetate/ethanol following standard procedures. The RNA concentration was quantified, diluted to 100 ng/μl and stored at -80°C.

*Production of Cas9 mRNA*

*Cas9* mRNA was transcribed from the pMLM3613 vector (Addgene plasmid #42251, Keith Joung) [3] using a mMESSAGE mMACHINER T7 Ultra Kit (Life Technologies). This vector has a unique *PmeI* restriction site positioned 3' at the end of the Cas9 coding sequence. The *Cas9* mRNA was DNase treated and precipitated with ammonium acetate/ethanol following standard procedures. The RNA concentration was quantified, diluted to 500 ng/μl and stored at -80°C.

*Microinjection of sgRNA/Cas9 in one-cell stage embryos*

Embryos were collected by natural mating of pairs of WT zebrafish. 200-250 one cell stage embryos were co-injected with approximately 1 nl total volume of Cas9-encoding mRNA (250 ng/μl) and sgRNA (25 ng/μl) each. The day after microinjection, ten normally developed 24 hpf embryos were selected to check targeting efficiency using T7EI assay

*Embryonic genomic DNA extraction and T7EI assay*

Genomic DNA was extracted from ten 24 hpf injected embryos and every individual sample was mixed with WT genomic DNA. For genomic DNA extraction a single embryo was placed into a microcentrifuge tube containing 20 μl of base solution (25 mM NaOH and 0.2 mM EDTA) and heated to 95°Cfor 30 minutes. The tube was then cooled to 4°C and 20 μl of neutralization buffer (40mM Tris-HCl pH 5.0) were added. The sample was centrifuged and the supernatant was used directly as template for PCR. We checked the occurrence of mutations and the efficiency of the method by using the T7EI assay. Once the positive outcome of the procedure was confirmed, the rest of the injected embryos were raised to adulthood to form the F0 generation.

*Screening by amplicon sequencing*

In order to identify a founder carrying a mutation of interest that could be transmitted through the germline, we outcrossed a single F0 fish with a WT fish and collected groups of five embryos produced by each breeding pair. After the extraction of genomic DNA from these pools of embryos (40 μl of base solution were used to cover the embryos and 40 μl of neutralization

buffer were subsequently added, as previously described), the samples were analysed by MiSeq Illumina sequencing. Specific primers containing a partial Illumina adaptor sequence were designed (Fw, 5'-<u>TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG</u>CATCTGCTGGACAAGTTTCTGA-3'; Rv, 5'-<u>GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG</u>ATAATTGGTGTCCGGGTCAAAC-3') and a region of about 230 bp surrounding the CRISPR target site was amplified by PCR. The quality and size of the amplicons produced by PCR were checked by running 5 µl of each PCR product on an agarose gel. In case of good quality and right size the rest of the samples were purified using a PCR & DNA Cleanup Kit (5 µg) (MonarchR, New England Biolabs). The DNA concentrations were quantified using a Qubit spectrophotometer and the amplicons were diluted to 15-25 ng/µl prior to analysis. Samples were sequenced using a MiSeq Illumina platform in collaboration with Dr Jason Rihel, University College London, UK. The read out of this analysis were FastQ data files, which were analysed using Geneious software. Once a F0 fish carrying an interesting indel transmitted to the germline was identified, the F1 embryos born from the cross between the selected F0 founder and WT were raised to adulthood. F1 fish were then tested with the T7EI assay and genotyped. Only F1 carrying the mutation were kept. We tested 40 F1 fish with the T7EI assay and found five fish carrying the 7 bp mutation (Supplementary Figure 4B).

*Genotyping*

Genotyping was performed by PCR reaction. To discriminate between WT, *ywhaz*$^{+/-}$ and *ywhaz*$^{-/-}$ fish, we designed specific primers that recognise the presence or the absence of the $\Delta 7$ allele. We also designed a pair of control primers to confirm the efficiency of the PCR reaction (Supplementary Table 1). To isolate genomic DNA for the PCR reaction from the fin, adult zebrafish were previously anesthetized in 0.02% tricaine methanesulfonate.

*Checking for the presence of off-target cleavage sites and generation of a stable mutant line*

To check the absence of any off-target cleavage before crossing F1, we used the CHOPCHOP web tool. We found two potential off-target sites for our sgRNA (Supplementary Table 2), that were amplified by PCR in our selected F1 fish using the following primers: 5'-CCTCTTGTCAGTGGCGTACA -3' (Fw) and 5'-TTCTCGGAGGACTCAACCAC -3' (Rv), for the off-target site present in *ywhag2*, and 5'- GCAGAGGAATTACGGATGGA -3' (Fw) and 5'-CGCGTTTATCCTGAGCTTTC -3' (Rv) for the one in *tspan9b*. PCR products were purified using a PCR & DNA Cleanup Kit (5 µg) (MonarchR, New England Biolabs), and analysed by Sanger sequencing. No extra mutations were found in the off-target areas of the gene (data not shown) and the selected F1 fish were in-crossed to obtain a final F2 stable mutant line homozygous for the mutation of interest.

**Gene expression analysis using RT-qPCR**

Total RNA was using the GenEluteTM Mammalian Total RNA Miniprep Kit (Sigma-Aldrich). Samples were DNase treated with TURBOTM DNase (ThermoFisher Scientific) to remove any genomic DNA contamination and checked for degradation. First strand cDNA was synthesised from 0.25 µg of tRNA using the RevertAid First Strand cDNA Synthesis Kit (ThermoFisher Scientific).

RT-qPCR was performed on 10 brains per genotype (WT and *ywhaz*$^{-/-}$) with three replicates for each brain using a CFX ConnectTM Real-Time System machine (BioRad Laboratories), the SensiFASTTM SYBR No-ROX Mix (Bioline), and the following primers: Fw 5'-GAGTACCGTGAGAAGATCGAAGC -3' and Rv 5'- CGGATCAGAAACTTGTCCAGCAG -3' (NCBI Reference Sequence: NM_212757). A melting curve step (50–95°C) was performed to verify that only single products had been amplified. No-template and no-reverse transcriptase controls were also performed for each primer pair and cDNA, respectively. To assess RT-qPCR efficiency, a 2-fold dilution series of cDNA template were processed. For normalization, expression levels of ribosomal protein L13a (*rpl13*) and elongation factor 1a (*elf1a*) were used as reference. The relative expression of the genes and the fold change were calculated using the $2^{-ddCT}$ comparative method [4,5].

**In situ hybridization (ISH)**

*Preparation of ywhaz mRNA probe*

Total RNA (tRNA) was extracted from whole frozen adult WT zebrafish brains using the TRIzol reagent. Complementary DNA (cDNA) was synthesized from 1 µg tRNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific). The *ywhaz* mRNA sequence was taken from the National Center for Biotechnology Information (NCBI) web site (NCBI Reference Sequence: NM_212757.2). Primers to amplify the open reading frame (ORF) of the gene were designed as follow: forward (Fw) primer 5'- AACCTGCTCTCTGTGGCCTA -3' and reverse (Rv) primer 5'-GCTCAGAAATGGCATCATCA -3'. The 481 bp *ywhaz* amplicon was generated by PCR reaction and cloned into a plasmid using the StrataClone PCR cloning Kit (Agilent). The plasmids were collected and purified using GeneJET Plasmid Maxiprep Kit (Thermo Scientific) and the product was sequenced by GATC Biotech to check the orientation of the insert in the plasmid and the identity of the sequence. The StrataClone PCR Cloning Vector pSCA-Amp/Kan containing the *ywhaz* insert was linearized with *NotI* restriction enzyme and *ywhaz* DIG-antisense RNA probe

was then generated by in vitro transcription. The product was DNase treated and cleaned using sodium acetate/ethanol precipitation. The final *ywhaz* probe was stored at -20°C.

*Preparation of the samples for ISH*

Embryos were treated with 1-phenyl 2-thiourea (PTU) at 24 hours post fertilization (hpf) to prevent pigmentation. Embryos, larvae and dissected brains from adult fish were fixed overnight at 4°C in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Specimens were then dehydrated with a gradient of methanol/PBS (25%, 50%, 75% and 100% methanol) before being stored for at least one hour and up to several months at -20°C.

*ISH protocol*

*First day of ISH*. Samples were rehydrated with a gradient of methanol/PBS (75%, 50%, 25% and 0% methanol) and then digested with proteinase K (10 µg/ml in PBS) at room temperature (30 minutes for a whole adult brain and 9 dpf embryos, 25 minutes for 6 dpf embryos, 15 minutes for 3 dpf embryos and 10 minutes for 2 dpf embryos). Samples were then fixed in 4% PFA for 20 minutes and rinsed in phosphate-buffered saline + 0.1% Tween-20 (PBT). Samples were prehybridized at 68°C for at least 2 hours in 300 µl of HYB+ buffer (65% formamide, 5X saline-sodium citrate (SSC) buffer, 50 µg/ml heparin, 0.5 mg/ml torula RNA, 0.1% Tween-20, 9.2 mM citric acid, pH 6.0). HYB+ was then replaced with fresh HYB+ buffer containing the DIG-labelled probe (5 ng/µl) and incubated overnight at 68°C.

*Second day of ISH*. The HYB+/probe mix was removed and stored at -20°C for future use. Samples were washed with a gradient of HYB+/2X SSC (75%, 50%, 25% and 0% HYB) for 10 minutes each, and then twice with 0.05X SSC for 30 minutes each. For ISH on sections, adult brains were fixed for 20 min with 4% PFA and embedded in 3% agarose dissolved in water. Samples were sectioned at 100 µm using a vibratome and sections were collected in PBS. Specimens were blocked for one hour at room temperature (RT) in blocking solution (2% normal goat serum, 2 mg/ml bovine serum albumin in PBT) and then incubated overnight with anti-DIG-AP antibody (1:4000 dilution in blocking solution).

*Third day of ISH*. Samples were washed several times in PBT and then three times for 10 minutes each in Xpho solution (100 mM Tris-HCl pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl and 0.1% Tween-20). Xpho solution was replaced with NBT/BCIP solution (225 µg/ml of NBT and 175µg/ml of BCIP in Xpho) and the specimens were incubated in the dark to develop the stain. Samples were monitored with a dissecting microscope every 30 minutes. The reaction was stopped by several washes in PBS and were fixed in 4% PFA for 20 minutes. Embryos were stored in 80% glycerol and 20% PBT at 4°C, whereas sections were mounted on slides and covered with Mowiol

solution. The NBT/BCIP signal was imaged using a GX microscope, a CMEX 5.0 camera and Image focus 4 software.

**Immunohistochemistry (IHC)**

Adult brains were rehydrated with a gradient of methanol/PBS (75%, 50%, 25%, 0% methanol) and then digested with proteinase K (10 μg/ml in PBS) at room temperature for 30 minutes followed by post-fixation in 4% PFA for 20 minutes. Samples were then rinsed in PBT. Specimens were embedded in 3% agarose dissolved in water and sectioned at 100 μm using a vibratome. Sections from *Tg*(*olig2*:*egfp*)[vu12] were blocked for 1 hour at RT with blocking solution and then incubated overnight in primary antibody. The day after, samples were washed several times in PBT, blocked with blocking solution for an hour at RT and incubated in secondary antibody diluted in blocking solution for 1 hour at RT. For *ywhaz-/-* and *Tg*(*aldoca*:*gap43-Venus*)[rk22] the Vectastain universal *Elite* ABC Kit (Vector, #PK-6200) was used. Therefore, the sections were blocked for 1 hour at RT with the blocking serum. The blocking serum consists of 1:10 normal horse serum in blocking diluent, which is 1%bovine serum albumin (BSA) in PBT. The specimens were then exposed overnight to primary antibody, followed by 1 hour incubation in secondary antibody, and 3 hours' incubation with Vectastain *Elite* ABC Reagent solution. After several washes in PBT, peroxidase activity was detected using 3,3'-Diaminobenzidine (DAB, Sigma, #D4293) according to the manufacturer's instruction. For double in situ hybridization/antibody labelling, the ISH was performed first (with NBT/BCIP staining) followed by the IHC (with DAB staining).

**High precision liquid chromatography (HPLC) analysis of monoamines and metabolites**

Samples were weighed, homogenised in 100 μl ice-cold 0.1 N perchloric acid using a pellet pestle (Sigma, #Z359971) and centrifuged at 12,000 rcf for 15 min at 4°C. The supernatant was collected and stored at -80°C until use.

Samples (15 μl) were then injected for HPLC analysis using a Spark Triathlon refrigerated autosampler and separation was achieved on a Luna C18(2), 5 μm, 100 Å, 100 x 1 mm (Phenomenex Ltd) reverse phase column. The mobile phase (75 mM sodium dihydrogen phosphate, 1 mM EDTA, 0.6 mM octane sulphonic acid (OSA) in deionised water containing 5% methanol, pH 3.7) was delivered by a Rheos 4000 pump (Presearch, UK). Electrochemical detection was performed using a glass carbon working electrode set at 700 mV relative to an

Ag/AgCl reference electrode, using an Antec Intro detector incorporating a low volume (VT-03) flow cell (Antec, Netherlands).

Samples were quantified by comparison with standard solutions of known concentrations of monoamines and metabolites using Chrom Perfect data analysis software (Justice Laboratories, NJ). Each sample was run in duplicate and the mean content of monoamine and neurotransmitter for sample was calculated and normalised to the weight of tissue. Results are expressed as picomoles per milligram of brain tissue.

**Whole-brain imaging**

*Transgenic zebrafish lines for brain imaging experiments*

*Tg(elavl3:GCaMP6s)* transgenic zebrafish larvae in the albino background [6,7] were used for brain imaging experiments as the WT group. To obtain a transgenic line expressing GCaMP6s pan-neuronally, with the albino background and not expressing *ywhaz*, we performed several crosses between the *Tg(elavl3:GCaMP6s)* and *ywhaz*$^{-/-}$ lines. First, F0 albino *Tg(elavl3:GCaMP6s)* fish were crossed with F0 *ywhaz*$^{-/-}$ zebrafish. F1 fish obtained from this breeding were crossed with albino *Tg(elavl3:GCaMP6s)* and F2 fish obtained from this breeding were genotyped to select only albino *Tg(elavl3:GCaMP6s) ywhaz*$^{+/-}$ fish. Protocol used for *ywhaz* genotyping is described above. To select *Tg(elavl3:GCaMP6s)* fish we performed a specific PCR that amplifies selectively a GCaMP6s fragment and fluorescence of larvae was confirmed with a fluorescence microscope. Selected F2 fish were in-crossed and F3 fish were genotyped to select albino *Tg(elavl3:GCaMP6s) ywhaz*$^{-/-}$ fish. F3 fish were then in-crossed to obtain a stable transgenic mutant line.

*Light-sheet microscope recordings*

Whole-brain imaging was performed using a custom-build Light Sheet Fluorescence Microscopy (LSFM) system in a configuration known as inverted Selective Plane Illumination Microscopy (iSPIM). To generate the light sheet, the laser beam from an ArKr laser (Innova 70C Spectrum, Coherent) passes through an acousto-optic tuneable filter (AOTFnC-400.650-TN, AAOptoelectronics) to select the desired wavelength (488nm) and to control the beam power. The beam is expanded with a Galilean telescope, until this moment the beam is radially symmetrical and, from this moment on, the illumination arm will be explained in the two axis. In the "x axis", the beam is cippling by a slit and passes through the flat part of a cylindrical lens, it is reflected by a galvanometric mirror, which only reduces the beam size and changes its direction, and focused with a scan lens. In the "y axis", the beam passes through the cylindrical

lens and is focused on the galvanometric mirror, which moves the beam on the "y axis". Then, the beam passes through a scan lens, which collimates the beam and translates the angular movement from the galvo mirror to translation movement in the beam to generate the light sheet. Finally, the light sheet passes through a relay lens, formed by the tube lens and the microscope objective, and displaces it perpendicularly to its propagation axis.

The detection path is composed of a 20X Objective (NA=1, WD=2mm, Olympus), a GFP filter (Semrock, FF01-525/45-25), two 200 mm tube lens, an electrically tunable lens (Optotune, EL-16-40-TC-VIS-20D) and an achromatic 100 mm lens, fig 3. The total magnification is 11.1x with and a field of view of 600μm and a pixel size of 0.5850μm using a Hamamatsu orca 4 .v3 camera. The ETL function enables to change the detection plane synchronously with the light sheet, which compensates the illumination displacement and generates focused images in the plane of the camera.

The chamber with the Fluorinated ethylene propylene (FEP) tube (ID=0.7 mm OD=1mm) containing the zebrafish embryo was mounted into the setup, on top of a platform connected to a xyz motorized stage. The FEP tube capillary was connected, using a screw, to a stepper motor (L4018S1204-M6, Nanotec) for sample rotation. The x, y, z, position of the detection objective and the sample rotation were manually adjusted. A region of interest (ROI) was selected when the whole brain of the larva was visible and the hemispheres were symmetrical (indicating a horizontal position of the larva). Once an ROI was selected, the GFP signal was recorded for 20 min at a scanning speed of 1 volumes/second and 60 planes/volume and 3ms exposure time per frames, every volume had a depth of 300um, with a voxel size of 0.585um x 0.585um x 5um (x, y, z).

**Behavioural tests**

All behavioural experiments were completed between 10:00 and 17:00 and recorded using FlyCapture2 2.5.2.3 software and a digital camera from Point Grey Research. Fish were left for 30 minutes to habituate to the testing room before the experiment. Adult fish were tested for social behaviour, anxiety, locomotion and aggression, and juvenile fish were tested for social behaviour. For adult fish: n = 12/genotype (1st batch), n=14 WT and 13 *ywhaz*[-/-] (2nd batch) and n = 13/genotype (3rd batch). For juvenile fish: n = 10/genotype.

*Visually-mediated social preference test (VMSP)*

The experiment was performed in two steps as described in Carreño Gutierrez et al., 2019 [8]. We used a mixture of size-matched males and females as strangers since they can attract both

male and female zebrafish [9]. Time spent in different zones of the tank was quantified using Noldus Ethovision XT software. The same procedure was used to test juvenile fish using a similar tank with a 1:3 reduction in size. The effect of novel non-social stimuli was also tested using marbles instead of the second group of unfamiliar fish in the second step of the test.

*Shoaling test*

The shoaling experiment was performed following the protocol from Parker et al., 2013 [10]. We used ViewPoint Life Sciences to measure the nearest neighbour distance (NDD) and the inter-individual distance (IID). We also virtually divided the tank into eight equal sections and calculated the cluster score [10]. The same procedure was used to test juvenile fish using a similar tank with a 1:3 reduction in size. Adult fish: n = 2 groups of 5 fish/genotype. Juvenile fish: n = 4 groups of 5 fish/genotype.

*Novel tank test (NTT)*

The NTT was performed in a standard 1.5 L trapezoid tank [11] and fish were recorded for 5 minutes. Noldus Ethovision XT software was used to measure the amount of time spent in the bottom (geotaxis) and top third of the tank, the time spent freezing, the total distance swum, the velocity and the absolute angular velocity.

*Open field test*

The open field test was performed in a large open tank (20.5 cm x 29 cm x 19 cm) and fish were recorded from above for 5 minutes. We used Noldus Ethovision XT software to quantify the duration of thigmotaxis (time spent swimming at a distance of 2 cm or less from the walls), the time spent in the periphery and in the centre, the time spent freezing, the distance swum and the velocity.

*Aggression test*

Aggression was measured using the mirror-induced stimulation protocol [12,13]. Fish were placed in holding tanks with walls covered in white paper on the night before the experiment to habituate them to the setup. Single fish were recorded for 5 minutes. The time spent in agonistic interaction was manually quantified using LabWatcher software from ViewPoint Life Sciences.

*Black and white test*

The black and white test was performed in a rectangular tank (24 cm x 12 cm) divided into two equal areas, a black area and a white area. Fish were placed in the centre of the tank and recorded for 5 minutes. The time spent in each area was manually quantified.

## 2. SUPPLEMENTARY TABLES

**Supplementary Table 1.** *ywhaz* genotyping primers designed to discriminate between WT, *ywhaz* +/- and *ywhaz* -/- fish.

| Specific primers of Δ7 allele | *ywhaz* forward (5'3') | TGACCTGGTTTCTGAGCTGA |
|---|---|---|
| | Mutant reverse (5'-3') | TGTAGCGACTTCTAGCGGT |
| Specific primers of WT allele | *ywhaz* forward (5'3') | TGACCTGGTTTCTGAGCTGA |
| | WT reverse (5'-3') | TAGCGACTTCTGCCAGGTAG |
| Internal control primers | Control forward (5'-3') | TGTACAAGTGCAGAAACCCAC |
| | Control reverse (5'-3') | TATCCGAATCAAGGCCAGGA |

**Supplementary Table 2.** Off-target sequences determined by CHOPCHOP.

| Number of mismatches | Off-target sequences | Alignment between CRISPR target sequence and off-target sequences | Position | Gene name |
|---|---|---|---|---|
| 2 | GGGAGATTATTACCGCTACCTGG |  | Exonic | *ywhag2* |
| 4 | GGGGGCCTATTACGGCTCCCGGG |  | Intronic | *tspan9b* |

**Supplementary table 3.** Pimary and secondary antibody concentration used for the IHC assays.

| Primary antibody | Secondary antibody | Genotype |
|---|---|---|
| **Anti-GFP** (Amsbio, #TP401) dilution 1:500 | Peroxidase Anti-rabbit igG (H+L) (Vector, #PI-1000), dilution 1:1000 | *Tg*(*olig2:egfp*)[vu12] |
| **Anti-parvalbumin** (Millipore, #MAB1572) dilution 1:1000 | Biotinylated Universal antibody anti-mouse and rabbit IgG (H+L) (Vectastain universal *Elite* ABC Kit, Vector) | *ywhaz*-/- |
| **Anti-GFP** (Amsbio, #TP401) dilution 1:500 | Biotinylated Universal antibody anti-mouse and rabbit IgG (H+L) (Vectastain universal *Elite* ABC Kit, Vector) | *Tg*(*aldoca:gap43-Venus*)[rk22] |

## 3. SUPPLEMENTARY FIGURES
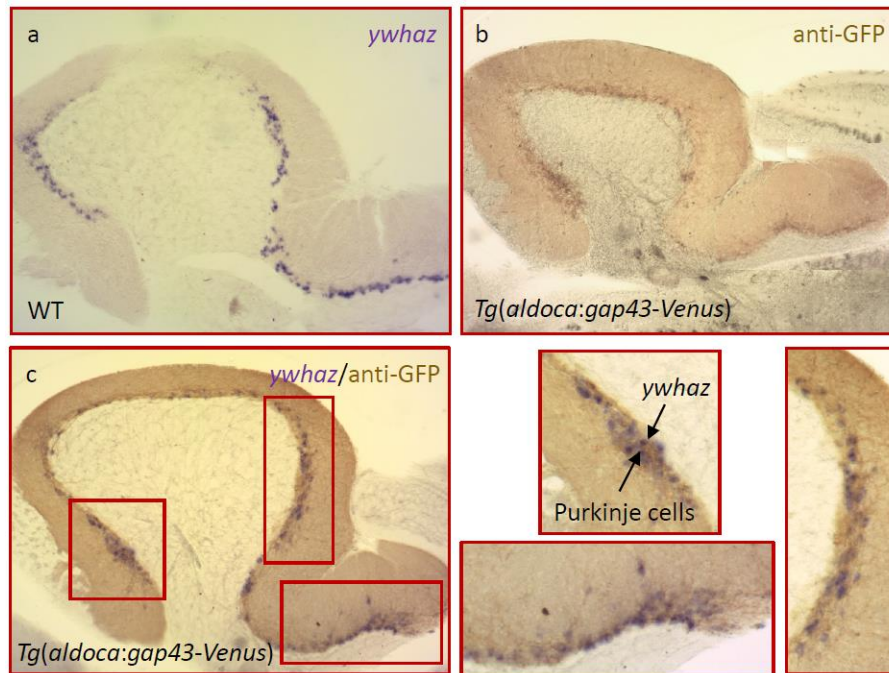


**Supplementary Figure 1. Co-staining of the *Tg(aldoca:gap43-Venus)* cerebellum with anti-GFP antibody and *ywhaz* riboprobe.** (a) *ywhaz* riboprobe staining (purple) performed by ISH on a sagittal section of WT cerebellum. (b) Anti-GFP antibody staining (brown) performed by IHC on a sagittal section of *Tg(aldoca:gap43-Venus)* cerebellum. (c) Co-staining of the *Tg(aldoca:gap43-Venus)* cerebellum with *ywhaz* riboprobe (purple), and anti-GFP antibody (brown) on a sagittal section of *Tg(aldoca:gap43-Venus)* cerebellum. The overlap between the two staining (black arrows) indicates that *ywhaz* is localised within Purkinje cells.
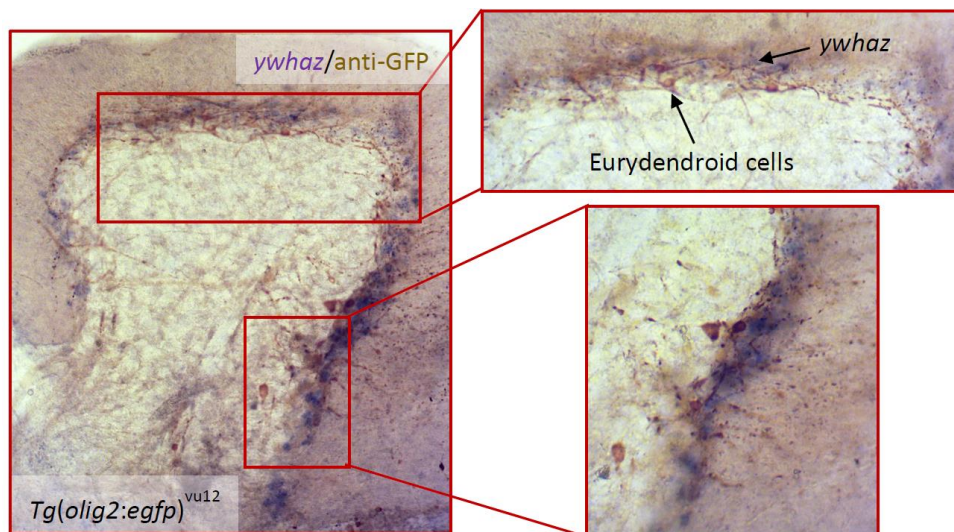


**Supplementary Figure 2. Co-staining of the *Tg(olig2:egfp)*[vu12] cerebellum with a *ywhaz* riboprobe (purple staining) and anti-GFP antibody (brown staining) in a sagittal section.** The absence of overlap between the two stainings shows that *ywhaz* is not localized in Eurydendroid cells.
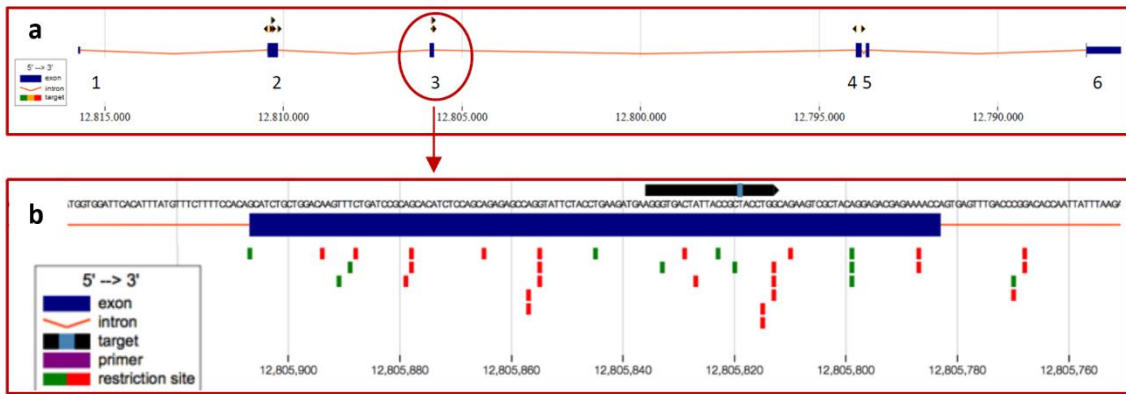
**Supplementary Figure 3. Design of target sequence and sgRNA for *ywhaz* using the CHOPCHOP web tool.** (a) All the possible target sequences in *ywhaz* (arrowhead) found by CHOPCHOP are shown above every exon (blue bars). Introns are represented as red lines in between exons. (b) The target sequence selected (black bar) within the third exon (blue bar) is represented here in greater details. Figures are taken and adapted from http://chopchop.cbu.uib.no.
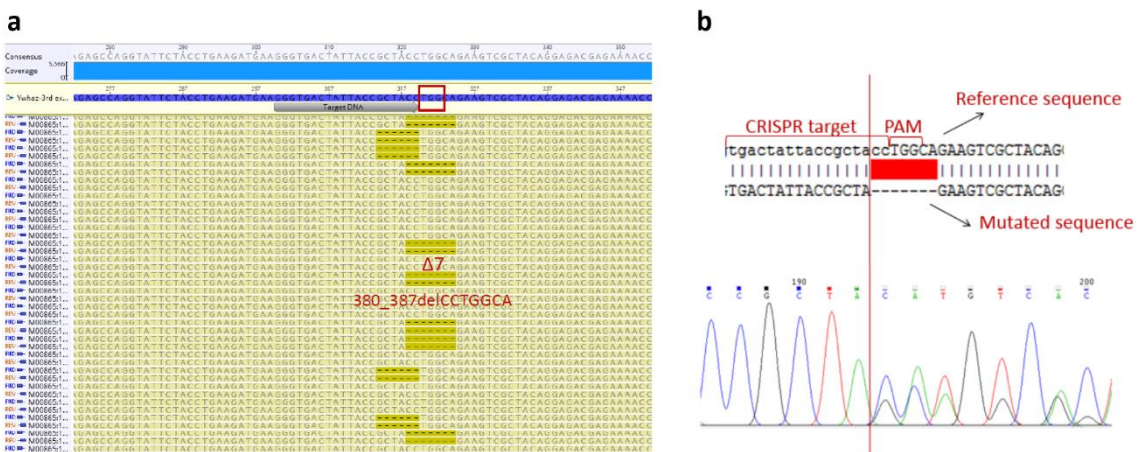


**Supplementary Figure 4. CRISPR/Cas9-mediated deletions of 6 and 7 bp in *ywhaz*.** (a) Read out of the MiSeq Illumina analyisis indicating the deletions of 6 and 7 bp in *ywhaz* caused by the microinjection of Cas9 mRNA and sgRNA. The grey arrow represents the *ywhaz* target DNA sequence, and the red rectangle the PAM sequence. MiSeq reads (yellow) are paired and aligned to the gene-specific reference sequence (purple). Dashed lines represent deletions. (b) Top, reference sequence containing the CRISPR target sequence and PAM motif within the third exon of *ywhaz* aligned with the correspondent sequence carrying the 7 bp deletion, indicated in red. Below, Sanger sequencing showing the frameshift introduced by the 7 bp deletion.
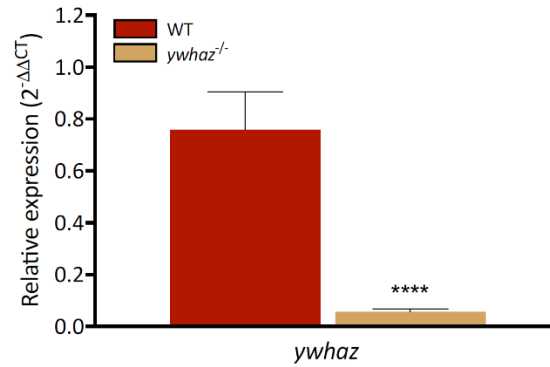
**Supplementary Figure 5. Effects of the selected 7bp deletion on *ywhaz* gene expression.** Relative expression profile of *ywhaz* normalised to the reference gene elongation factor 1a (*elf1a*). *ywhaz*[-/-] have a significantly decreased level of *ywhaz* expression compared to WT (p < 0.0001, Mann-Whitney unpaired t-test, n=10 WT, n=10 ywhaz[-/-]). **** p<0.0001. Mean ± SEM.
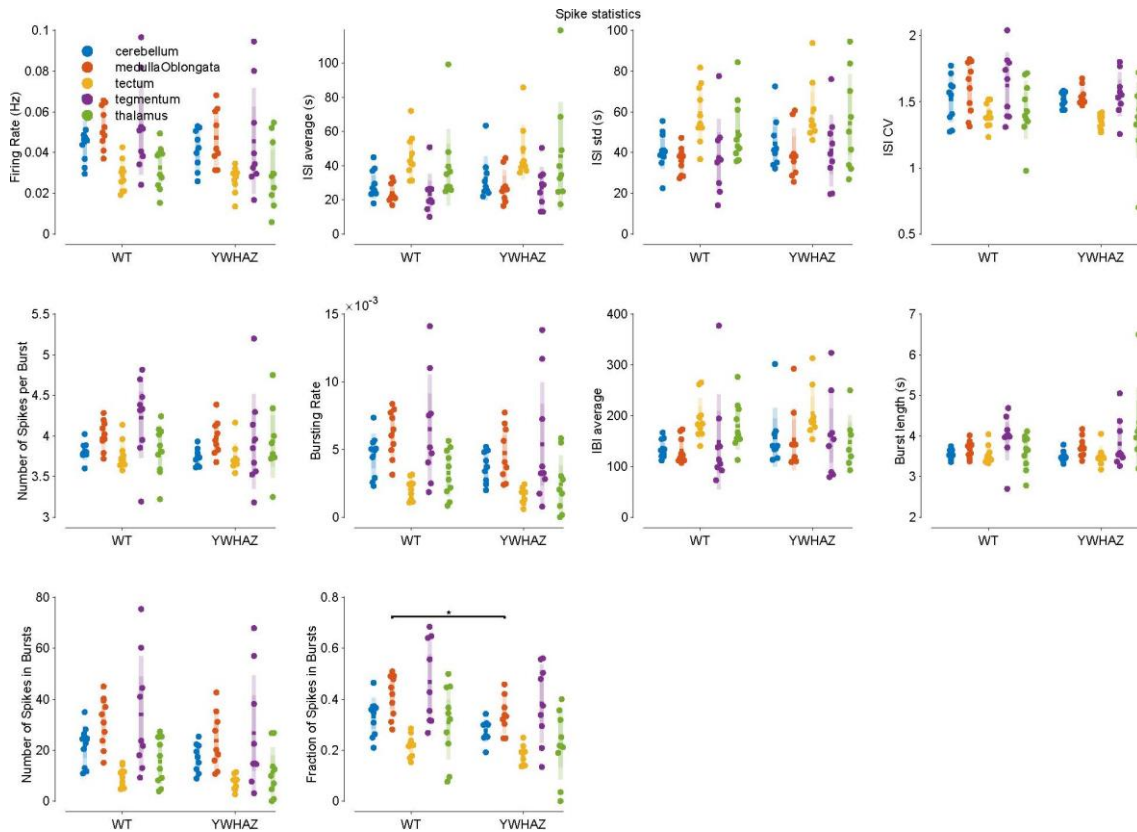


**Supplementary Figure 6. Single-cell activity analysis performed in the five defined brain areas.** n= 10 WT and 9 *ywhaz*[-/-]. Unpaired t-tests, each single point represents an individual, the central line represents the mean, the darker bar the 95% confidence interval and the lighter bar the standard deviation. YWHAZ, *ywhaz*[-/-] larvae; WT, wild-type larvae. * p < 0.05
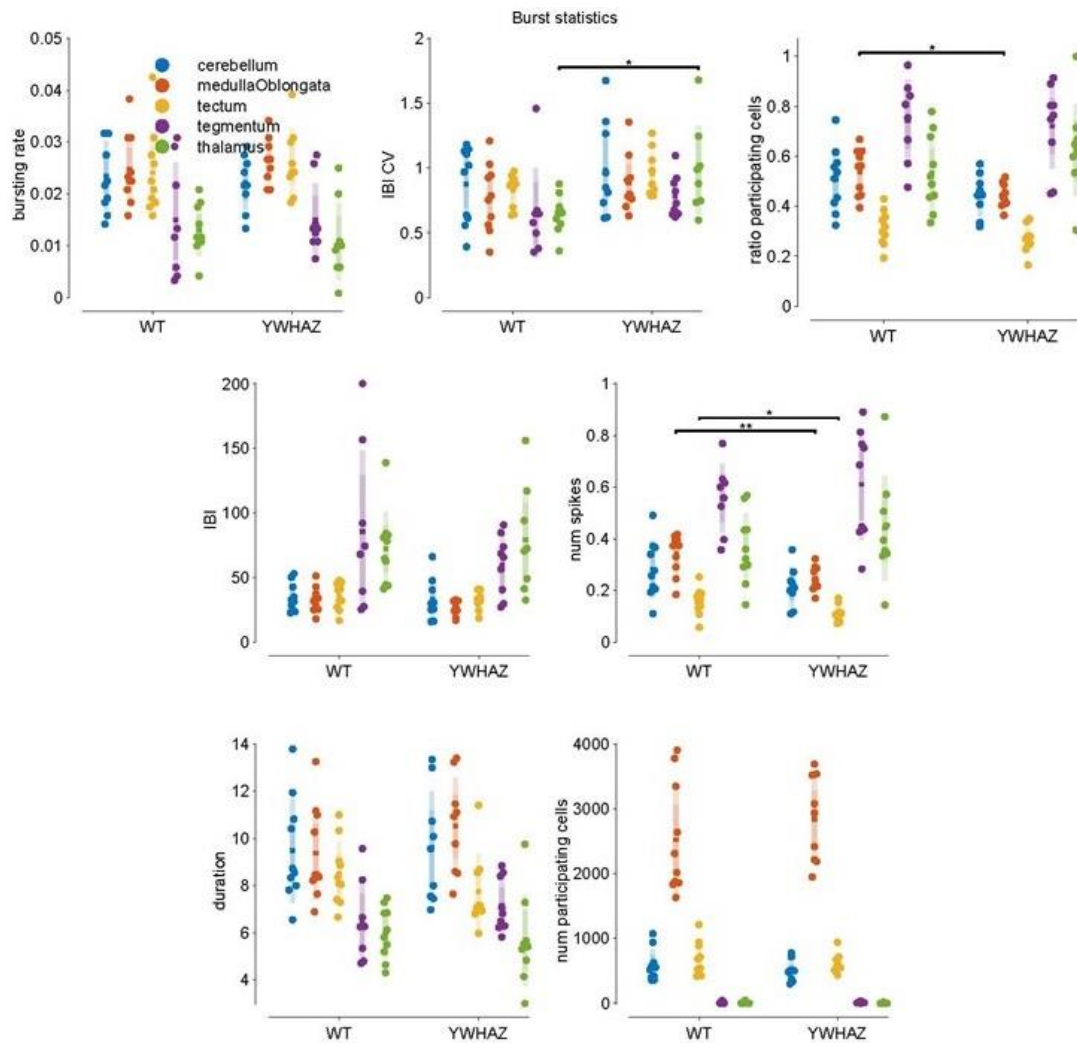
**Supplementary Figure 7. Collective bursts activity analysis performed in the five defined brain areas.** n= 10 WT and 9 *ywhaz*[-/-]. Unpaired t-tests, each single point represents an individual, the central line represents the mean, the darker bar the 95% confidence interval and the lighter bar the standard deviation. YWHAZ, *ywhaz*[-/-] larvae; WT, wild-type larvae. * p < 0.05, ** p < 0.01.
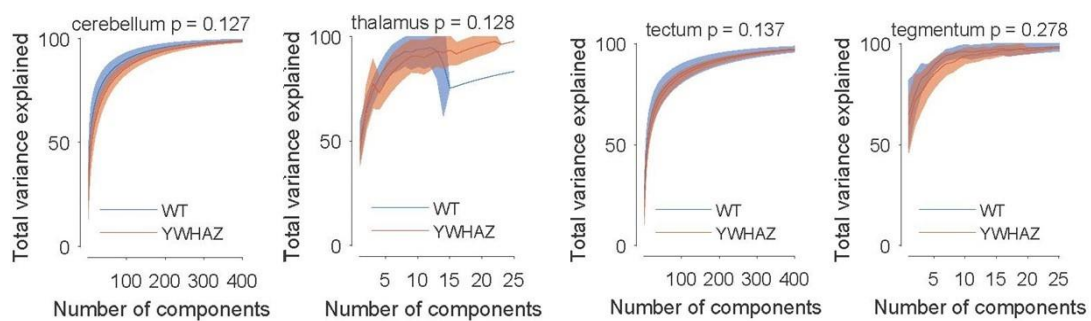


**Supplementary Figure 8. Principal component analysis of neuronal activity performed in cerebellum, thalamus, tectum and tegmentum.** n= 10 WT and 9 *ywhaz*[-/-]. The central line represents the mean and the coloured shadow represents the 95% confidence interval. YWHAZ, *ywhaz*[-/-] larvae; WT, wild-type larvae.
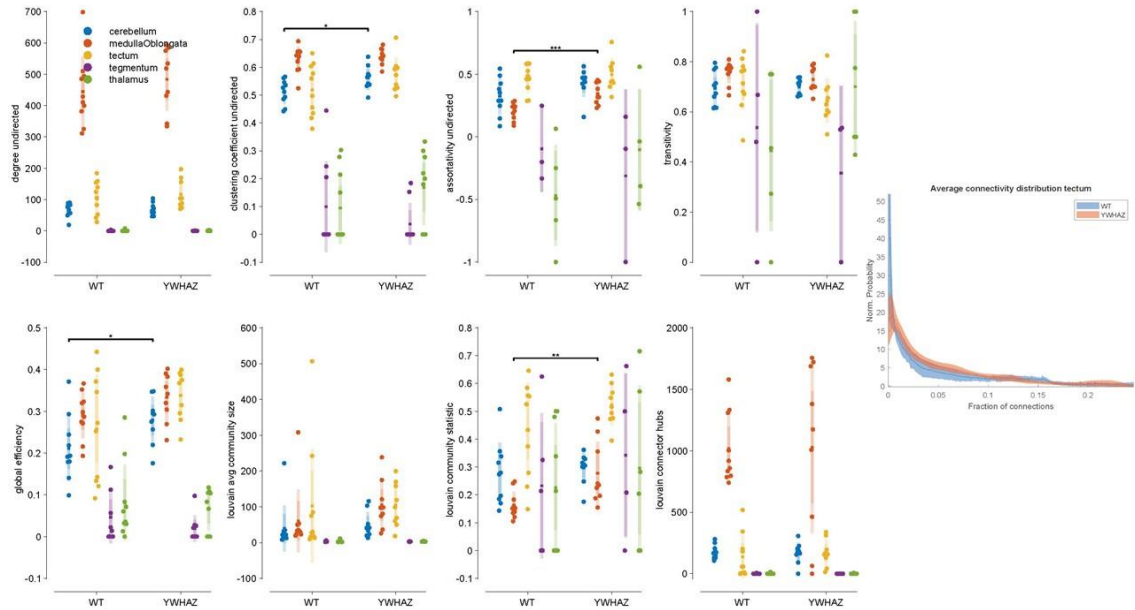
**Supplementary Figure 9. Connectivity analysis performed in the five defined brain areas and neuronal connectivity distribution in the optic tectum.** On the left, unpaired t-tests of connectivity parameters, each single point represents an individual, the central line represents the mean, the darker bar the 95% confidence interval and the lighter bar the standard deviation. On the right, connectivity distribution in the tectum of WT and *ywhaz*$^{-/-}$ larvae. The central line represents the mean and the coloured shadow represents the 95% confidence interval. n= 10 WT and 9 *ywhaz*$^{-/-}$. * p < 0.05, ** p < 0.01, *** p < 0.001.
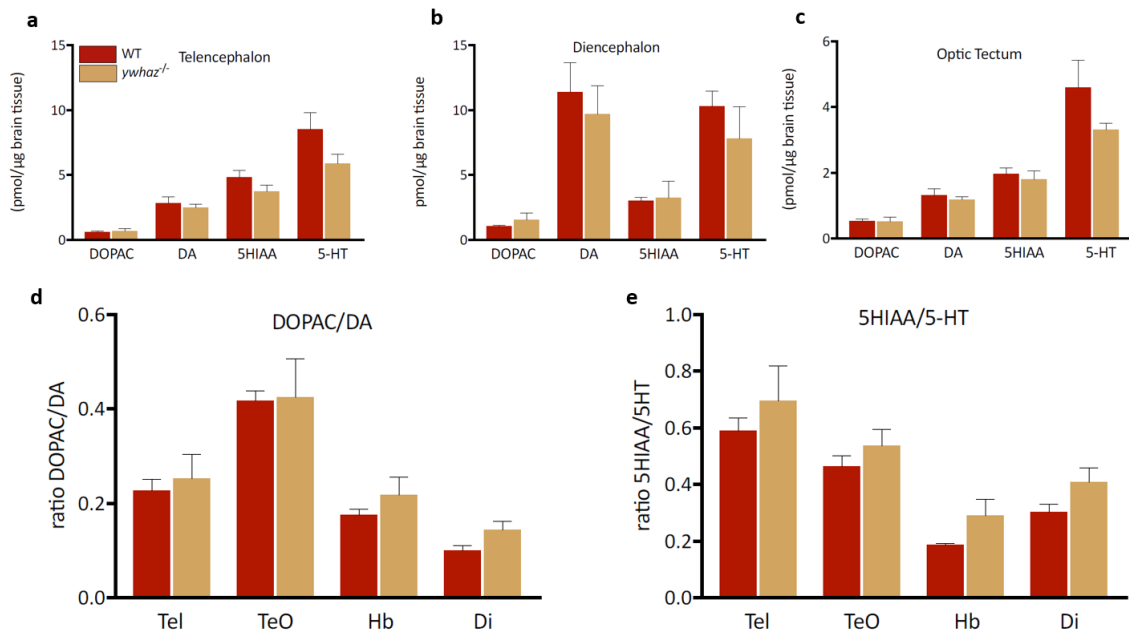
**Supplementary Figure 10. Analysis of the effect of *ywhaz* loss-of-function on monoamines and their metabolites levels in different areas of WT and *ywhaz*$^{-/-}$ adult brains by high precision liquid chromatography (HPLC).** No significant difference of monoamine levels was found in (a) telencephalon, (b) diencephalon and (c) optic tectum. There was also no significant difference in the breakdown of (d) DOPAC/DA and (e) 5HIAA/5-HT in any of the brain areas. Multiple t-tests with Holm-Sidak correction for multiple comparisons. Abbreviations: DA, dopamine; Di, diencephalon; DOPAC, 3,4-dihydroxyphenylacetic acid; Hb, hindbrain; Tel, telencephalon; TeO, optic tectum; 5-HIAA, 5-hydroxyindoleacetic acid; 5-HT, 5-hydroxytryptamine. n = 7 WT, n = 7 *ywhaz-/-*. Mean ± SEM.
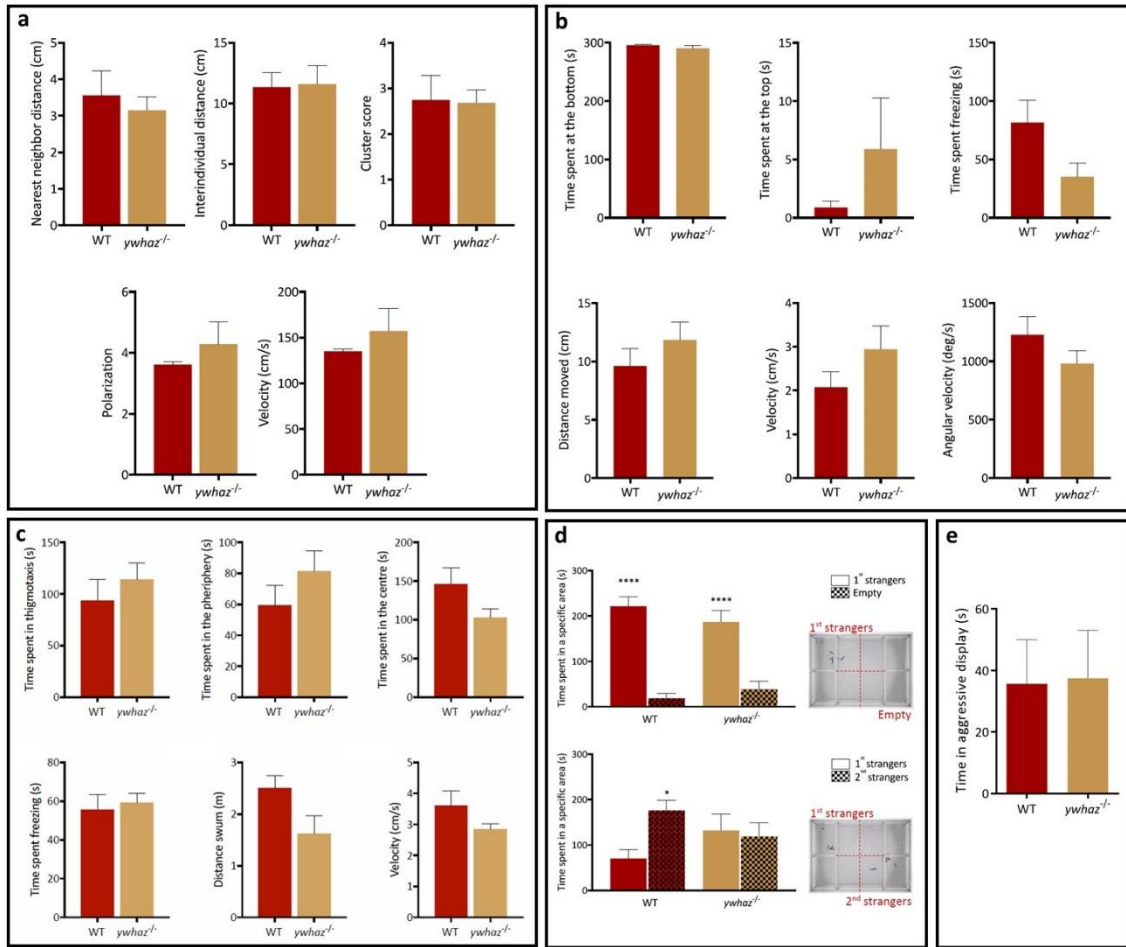
**Supplementary Figure 11. Behavioural tests performed in adult fish. (a) Shoaling behaviour.** Adult *ywhaz*[-/-]
display normal shoaling. Nearest neighbour distance (p = 0.66), inter-individual distance (p = 0.91), cluster score
(p = 0.93), polarization (p = 0.53) and velocity (p = 0.53). n = 2 groups of 5 wild-type (WT), n = 2 groups of 5
*ywhaz*[-/-]. **(b) Novel tank test.** *ywhaz-/-* exhibit normal anxiety-like behaviour. Time spent at the bottom (p =
0.41), at the top (p = 0.50), and freezing (p = 0.13) in a novel tank. n = 15 WT, n = 15 *ywhaz*[-/-]. Mann-Whitney U
test. Locomotion (p = 0.31), velocity (p = 0.19) and angular velocity (p = 0.22) in a novel tank. n = 15 WT, n = 15
*ywhaz*[-/-]. Unpaired t-test with Welch's correction. **(c) Open field test.** *ywhaz*[-/-] behave similarly to WT in the
open field test. Time at the side of the tank (p = 0.42), in the periphery (p = 0.24), in the centre of the tank (p =
0.19) and time spent freezing (p = 0.70). n = 15 WT, n = 15 *ywhaz*[-/-]. Unpaired t-test with Welch's correction.
Locomotion (p = 0.057). n = 15 WT, n = 15 *ywhaz*[-/-]. Mann-Whitney U test. Velocity (p = 0.16) is not affected in
the open field test. n = 15 WT, n = 15 *ywhaz*[-/-]. Unpaired t-test with Welch's correction. **(d) Visually-mediated
social preference test.** On the top, social preference step: both WT and ywhaz[-/-] show a significant preference
to spend time near a group of unfamiliar fish (1st strangers; p < 0.0001 for both WT and ywhaz[-/-], two-way
ANOVA with Tukey's post hoc comparisons, n = 12). On the bottom, preference for social novelty step: WT
switch preference and spend more time close to the second group of unfamiliar fish (2nd strangers; p = 0.048).
ywhaz[-/-] spend an equal amount of time near both groups of unfamiliar fish (1st and 2nd strangers; p = 0.98).
Two-way ANOVA with Tukey's post hoc comparisons, n = 12. **(e) Mirror-induced aggression test**. No difference
in aggression levels between WT and *ywhaz*[-/-] (p = 0.72). n = 15 WT, n = 15 *ywhaz*[-/-]. Mann-Whitney unpaired t-
test. Mean ± SEM.
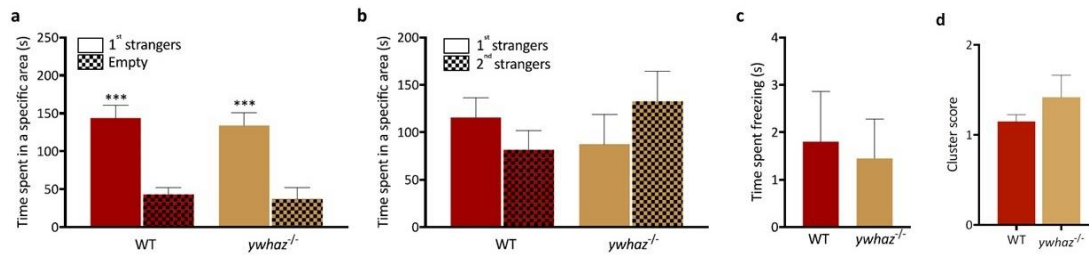
**Supplementary Figure 12. Visually-mediated social preference test and shoaling test with juvenile fish.**
(a) Social preference step. Similar to adults, both WT and *ywhaz*$^{-/-}$ juvenile fish show a significant preference to spend time near a group of unfamiliar fish (1$^{st}$ strangers; $p < 0.001$ for both WT and *ywhaz*$^{-/-}$). (b) Preference for social novelty step. WT and *ywhaz*$^{-/-}$ juveniles spend equal time close both groups of unfamiliar fish (2$^{nd}$ strangers; $p = 0.077$ and $p = 0.63$ respectively). Two-way ANOVA with Tukey's post hoc comparisons. n = 10. (c) Time spent freezing during the first two minutes after the addition of the second group of unfamiliar fish in the behavioural tank. Neither WT nor *ywhaz*$^{-/-}$ juveniles show any freezing reaction ($p > 0.99$). Mann-Whitney U test. n = 10. Mean ± SEM. (d) Shoaling behaviour is normal in juvenile fish, with equal cluster score between genotypes ($p = 0.33$). n = 4 groups of 5 WT, n = 4 groups of 5 *ywhaz*$^{-/-}$. Unpaired t-tests with Welch's correction. *** $p < 0.001$. Mean ± SEM.
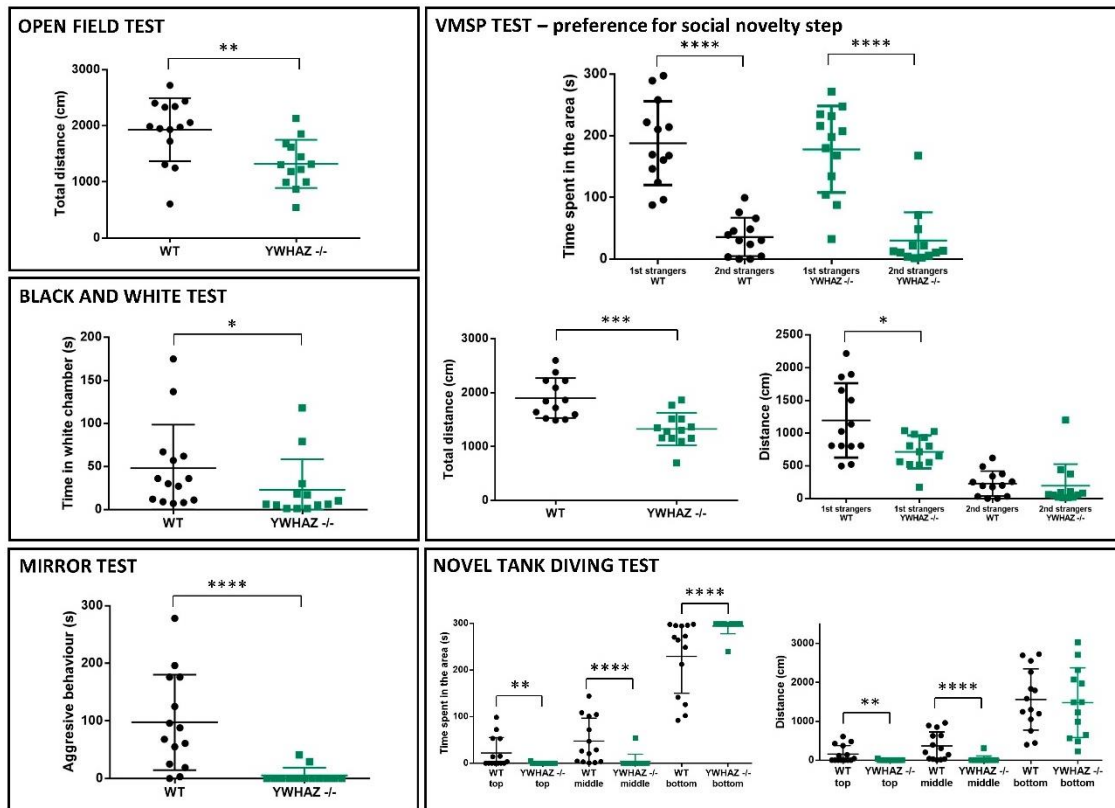
**Supplementary Figure 13. Behavioural tests repeated in a second batch of adult fish. Open field test.** Total distance travelled is significantly lower in *ywhaz⁻/⁻* compared to WT (p = 0.0040, Unpaired t-test with Welch's correction) due to an increased freezing observed in mutant fish **Visually-mediated social preference test.** *ywhaz⁻/⁻* behave similarly to WT in the preference for social novelty step and show preference for the first group of strangers (Time spent in the area: 1st strangers vs 2nd strangers; p < 0.0001; *ywhaz⁻/⁻* p < 0.0001; Two way ANOVA, no RM, followed by Sidak's post hoc test). However, *ywhaz⁻/⁻* present a freezing behaviour, reflected in the lower distance travelled by mutant fish (Total distance, p = 0.0002; Distance travelled in the 1st strangers area, p = 0.0126). Unpaired t-tests with Welch's correction. **Black and white test.** *ywhaz-/-* spend less time in the white chamber than WT fish (p= 0.0199, Mann-Whitney U test). Indeed, they spend more time freezing in the black chamber. **Mirror-induced aggression test**. *ywhaz⁻/⁻* spend less time performing an aggressive behaviour (p < 0.0001, Mann-Whitney U test), as they spend most of the time freezing. **Novel tank diving test.** *ywhaz⁻/⁻* exhibit a higher anxiety-like behaviour: they spend more time than WT fish at the bottom area (p < 0.0001), and less time at the middle (p < 0.0001) and at the top (p = 0.0.63) areas of the novel tank. *ywhaz⁻/⁻* travel less distance at the middle (p < 0.0001) and top (p = 0.0050) areas of the novel tank. Mann-Whitney U tests. For all the experiments, n= 14 WT, n = 13 *ywhaz⁻/⁻*. Mean ± SD. * p < 0.05; ** p < 0.01; *** p < 0.001; **** p<0.0001.
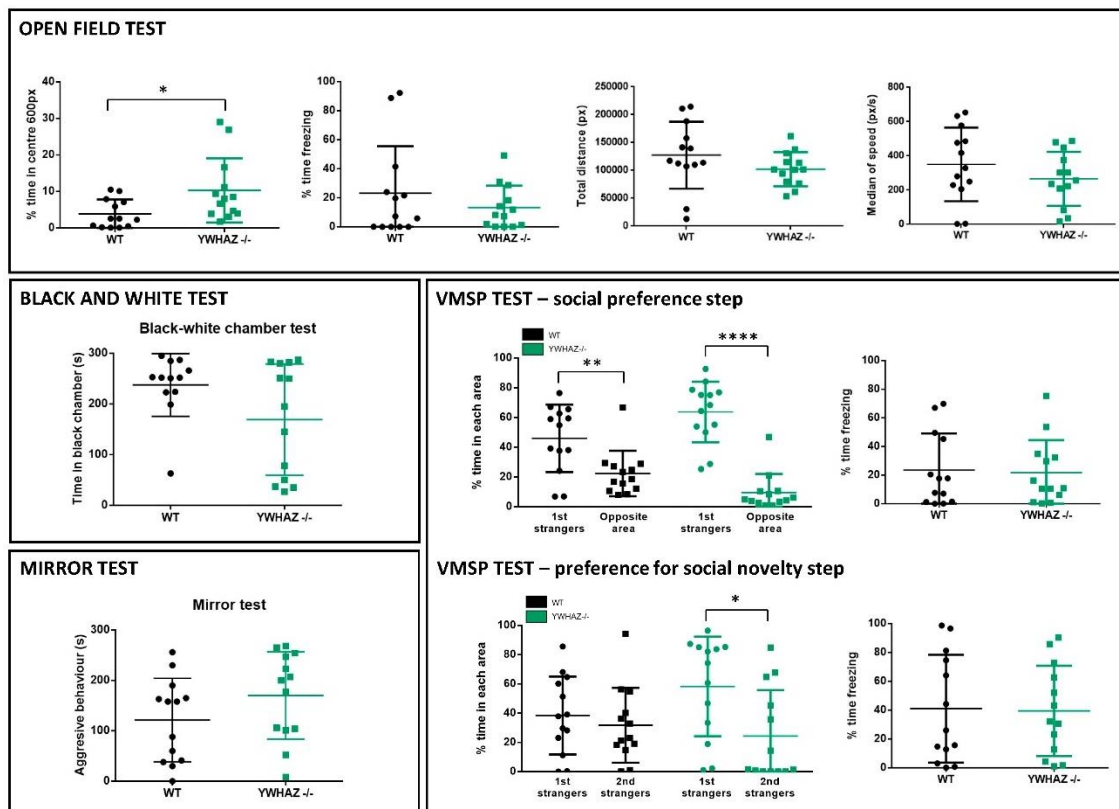
**Supplementary Figure 14. Behavioural tests repeated in a third batch of adult fish in a different setup.**
**Open field test.** *ywhaz*$^{-/-}$ behave similarly to WT in the open field test although they spend more time in the centre of the arena (p=0.016, Mann-Whitney U test). Time freezing (p = 0.83, Mann-Whitney U test), total distance travelled (p = 0.19, Unpaired t-test with Welch's correction) and median of speed (p = 0.27, Unpaired t-test with Welch's correction). **Black and white test.** Although no significant statistical differences (p = 0.12, Mann-Whitney U test), WT fish seem to be more anxious, as they spend almost the whole time in the black area while *ywhaz*-/- fish show no preference for any area. **Visually-mediated social preference test.** On the top, social preference step: both WT and ywhaz$^{-/-}$ show a significant preference to spend time near a group of unfamiliar fish (1$^{st}$ strangers vs opposite area: WT p = 0.0035, *ywhaz*$^{-/-}$ p < 0.0001; Two way ANOVA, no RM, followed by Sidak's post hoc test) and freeze a similar amount of time (p = 0.99, Mann-Whitney U test). On the bottom, preference for social novelty step: WT fish spend an equal amount of time near both groups of unfamiliar fish and ywhaz$^{-/-}$ spend more time close to the first group of strangers (1$^{st}$ strangers vs 2$^{nd}$ strangers: WT p = 0.814, *ywhaz*$^{-/-}$ p = 0.011; Two way ANOVA, no RM, followed by Sidak's post hoc test) **Mirror-induced aggression test**. No difference in aggression levels between WT and *ywhaz*$^{-/-}$ (p = 0.16, Unpaired t-test with Welch's correction). For all the experiments, n= 13 WT, n = 13 *ywhaz*$^{-/-}$. Mean ± SD. * p < 0.05; ** p < 0.01; **** p<0.0001.

## 4. REFERENCES

1.  Montague, T.G.; Cruz, J.M.; Gagnon, J.A.; Church, G.M.; Valen, E. CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **2014**, *42*, doi:10.1093/nar/gku410.

2.  Labun, K.; Montague, T.G.; Gagnon, J.A.; Thyme, S.B.; Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **2016**, *44*, W272–W276, doi:10.1093/nar/gkw398.

3.  Hwang, W.Y.; Fu, Y.; Reyon, D.; Maeder, M.L.; Tsai, S.Q.; Sander, J.D.; Peterson, R.T.; Yeh, J.R.J.; Joung, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 227–229, doi:10.1038/nbt.2501.

4.  Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **2001**, *25*, 402–408, doi:10.1006/meth.2001.1262.

5.  Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **2008**, *3*, 1101–1108, doi:10.1038/nprot.2008.73.

6.  Urasaki, A.; Morvan, G.; Kawakami, K. Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition. *Genetics* **2006**, *174*, 639–649, doi:10.1534/genetics.106.060244.

7.  Kawakami, K.; Takeda, H.; Kawakami, N.; Kobayashi, M.; Matsuda, N.; Mishina, M. A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev. Cell* **2004**, *7*, 133–144, doi:10.1016/j.devcel.2004.06.005.

8.  Carreño Gutiérrez, H.; Colanesi, S.; Cooper, B.; Reichmann, F.; Young, A.M.J.; Kelsh, R.N.; Norton, W.H.J. Endothelin neurotransmitter signalling controls zebrafish social behaviour. *Sci. Rep.* **2019**, *9*, doi:10.1038/s41598-019-39907-7.

9.  Ruhl, N.; McRobert, S.P.; Currie, W.J.S. Shoaling preferences and the effects of sex ratio on spawning and aggression in small laboratory populations of zebrafish (Danio rerio). *Lab Anim. (NY).* **2009**, *38*, 264–269, doi:10.1038/laban0809-264.

10. Parker, M.O.; Brock, A.J.; Millington, M.E.; Brennan, C.H. Behavioural phenotyping of casper mutant and 1-Pheny-2-Thiourea treated adult zebrafish. *Zebrafish* **2013**, *10*, 466–471, doi:10.1089/zeb.2013.0878.

11. Egan, R.J.; Bergner, C.L.; Hart, P.C.; Cachat, J.M.; Canavello, P.R.; Elegante, M.F.; Elkhayat, S.I.; Bartels, B.K.; Tien, A.K.; Tien, D.H.; et al. Understanding behavioral and physiological phenotypes of stress and anxiety in zebrafish. *Behav. Brain Res.* **2009**, *205*, 38–44, doi:10.1016/j.bbr.2009.06.022.

12. Gerlai, R.; Lahav, M.; Guo, S.; Rosenthal, A. Drinks like a fish: Zebra fish (Danio rerio) as a behavior genetic model to study alcohol effects. *Pharmacol. Biochem. Behav.* **2000**, *67*, 773–782, doi:10.1016/S0091-3057(00)00422-6.

13. Norton, W.H.J.; Stumpenhorst, K.; Faus-Kessler, T.; Folchert, A.; Rohner, N.; Harris, M.P.; Callebert, J.; Bally-Cuif, L. Modulation of fgfr1a signaling in zebrafish reveals a genetic basis for the aggression-boldness syndrome. *J. Neurosci.* **2011**, *31*, 13796–13807, doi:10.1523/JNEUROSCI.2892-11.2011.

# CHAPTER 2.

Exploring the contribution of *RBFOX1*

to ASD and other psychiatric disorders

# Article 3. One gene to rule them all: *RBFOX1* and mental disorders

**Summary in Spanish: "Un gen para gobernarlos a todos: *RBFOX1* y trastornos psiquiátricos"**

Estudios previos han sugerido que variantes comunes en el gen *RBFOX1* contribuyen al desarrollo de trastornos psiquiátricos mientras que variantes raras en este gen se han relacionado con el trastorno del espectro autista (TEA). En este estudio demostramos que existen asociaciones genéticas entre polimorfismos de nucleótido único situados en *RBFOX1* y el trastorno depresivo mayor, la tolerancia al riesgo y la esquizofrenia. Además, las variaciones en el número de copias que abarcan *RBFOX1* son más frecuentes en pacientes con TEA que en controles, pacientes que presentan una expresión disminuida de *RBFOX1* en corteza cerebral *postmortem*. Mediante estudios funcionales demostramos que portadores de un genotipo común de *RBFOX1* presentan una reactividad aumentada a estímulos emocionales, una menor eficiencia en el procesamiento prefrontal durante tareas que requieren control cognitivo y un aumento de la expresión del miedo tras un experimento de condicionamiento. Todo esto acompañado de un aumento del comportamiento de evasión. Además, un modelo murino genoanulado para *Rbfox1* presenta hiperactividad, estereotipias, problemas en el procesamiento del miedo, una disminución del interés social y una falta de agresividad, lo que valida este ratón genoanulado como modelo del TEA. En conjunto, estos resultados demuestran que variantes comunes en *RBFOX1* contribuyen a una amplia variedad de trastornos psiquiátricos mientras que variantes raras en *RBFOX1* con un mayor efecto están asociadas con un fenotipo autista. Por tanto, estos resultados evidencian a *RBFOX1* como gen de riesgo implicado en un amplio abanico de trastornos psiquiátricos

**Reference:**

O'Leary A, Fernàndez-Castillo N, Gan G, Antón-Galindo E, Cabana-Domínguez J, Yotova A, Kranz T, Grünewald L, Burguera D, Pané-Farré CA, Gerlach AL, Wittchen HU, Lang T, Alpers GW, Fydrich T, Ströhle A, Arolt V, Schweiger S, Winter J, Mota NR, Franke B, Harneit A, Schweiger JI, Schwarz K, Ma R, Chen J, Schwarz E, Tost H, Meyer-Lindenberg A, Erk S, Heinz A, Romanczuk-Seiferth N, Walter H, Witt S, Rietschel M, Noethen MM, Richter J, Yang Y, Kircher T, Hamm AO, Straube B, Lueken U, Weber H, Deckert J, Freudenberg F, Cormand B, Slattery DA, Reif A. One gene to rule them all: *RBFOX1* and mental disorders. To be submitted.

# *One gene to rule them all: RBFOX1 and mental disorders*

Aet O'Leary[1], Noèlia Fernàndez-Castillo[2,3,4], Gabriela Gan[5], Ester Antón-Galindo[2,3,4], Judit Cabana-Domínguez[2,3,4], Anna Yotova[1], Thorsten Kranz[1], Lena Grünewald[1], Demian Burguera[5], Christiane A. Pané-Farré[6], Alexander L. Gerlach[7], Hans-Ulrich Wittchen[8], Thomas Lang[9], Georg W. Alpers[10], Thomas Fydrich[11], Andreas Ströhle[12], Volker Arolt[13], Susann Schweiger[1], Jennifer Winter[1], Nina Roth Mota[14**¡Error! No se encuentra el origen de la referencia.**], Barbara Franke[14], Anais Harneit[15], Janina I. Schweiger[15], Kristina Schwarz[15], Ren Ma[15], Junfang Chen[15], Emanuel Schwarz[15], Heike Tost[15], Andreas Meyer-Lindenberg[15],Susanne Erk[16], Andreas Heinz[16], Nina Romanczuk-Seiferth[16], Henrik Walter[16],Stephanie Witt[17], Marcella Rietschel[17], Markus M. Noethen[18,19], Jan Richter[20], Yunbo Yang[21], Tilo Kircher[21], Alfons O. Hamm[20], Benjamin Straube[21], Ulrike Lueken[11], Heike Weber[1], Jürgen Deckert, Florian Freudenberg[1], Bru Cormand[2,3,4], David A. Slattery[1]**¡Error! No se encuentra el origen de la referencia.**, Andreas Reif[1]

[1]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Frankfurt, Germany

[1]Mainz University Medical Center, Institute of Human Genetics, Mainz, Germany

[2]Department de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, and Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain.

[3]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III (ISCIII), Madrid, Spain.

[4]Institut de Recerca Sant Joan de Déu (IRSJD), Esplugues de Llobregat, Barcelona, Catalonia, Spain

[5]Department of Zoology, Charles University, Vinicna 7, Prague, Czech Republic.

[6]Department of Psychology, Clinical Psychology, Experimental Psychopathology, and Psychotherapy, Philipps University Marburg, Marburg, Germany.

[7]Clinical Psychology and Psychotherapy, University of Cologne, Cologne, Germany.

[8]Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Dresden, Germany.

[9]Christoph-Dornier-Foundation for Clinical Psychology, Institute for Clinical Psychology Bremen, Bremen, Germany; Department for Psychology & Methods, Jacobs University Bremen, Bremen, Germany.

[10]Department of Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany.

[11]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany.

[12]Department of Psychiatry and Psychotherapy, Campus Charité Mitte, Charité-Universitätsmedizin Berlin, Berlin, Germany.

[13]Department of Psychiatry and Psychotherapy, University of Münster, Münster, Germany.

[14]Department of Human Genetics and Department of Psychiatry, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen 6500 HB, Netherlands

[15]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[16]Department of Psychiatry and Psychotherapy, Charité - University Medicine Berlin, Berlin, Germany

[17]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[18]Institute of Human Genetics, University of Bonn, Germany

[19]Department of Genomics, Life & Brain Center, University of Bonn, Germany

[20]Department of Biological and Clinical Psychology/Psychotherapy, University of Greifswald, Greifswald, Germany

[21]Department of Psychiatry and Psychotherapy, Phillipps-University Marburg, Marburg, Germany

[22]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University of Würzburg, Würzburg, Germany.

**Corresponding author:**

Dr. Aet O'Leary, Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt – Goethe University, Heinrich-Hoffmann-Str. 10, 60528 Frankfurt am Main, Germany; aet.oleary@kgu.de

## ABSTRACT

Common variation in the gene encoding the neuron-specific RNA splicing factor RNA Binding Fox-1 Homolog 1 (*RBFOX1*) has been suggested to contribute to several psychiatric conditions, while rare genetic variants have been found in autism spectrum disorders (ASD). Here we further mapped the genetic landscape of *RBFOX1* risk by demonstrating gene- and genome-wide association of *RBFOX1* single nucleotide variants with major depressive disorder, risk tolerance and schizophrenia. Moreover, copy number variants were more frequent in ASD cases than in controls. In line, *RBFOX1* expression was significantly decreased in *postmortem* cortical regions from patients suffering from ASD. Functional studies demonstrated that carriers of a common *RBFOX1* genotype display increased reactivity to emotional *stimuli*, less efficient prefrontal processing during cognitive control and enhanced fear expression after fear conditioning. This went along with increased avoidance behavior. These intermediate neural phenotypes were further specified by investigating neuronal-specific *Rbfox1* knockout (KO) mice. These are characterized by pronounced hyperactivity, stereotyped behaviour, impairments in fear acquisition and extinction, less social interest, and lack of aggression, and hence feature excellent construct and face validity as an animal model of ASD. The above evidence shows that common variants in *RBFOX1* lead to several psychiatric phenotypes, and rare genetic variations with large effect sizes are associated with increasing disorder severity. Therefore, our findings highlight *RBFOX1* as a prime risk gene across numerous mental disorders.

## MAIN ARTICLE

Mental disorders are characterized by substantial heritability of complex genetic architecture, with many variants of individual small effects interacting with rare variants of intermediate or large effects [1]. A high degree of comorbidity and shared heritability, along with evidence of substantial genetic correlations among psychiatric diseases, point to a role for pleiotropic effects. Also, genetic pleiotropy takes place in that loss-of-function variants cause rare severe genetic syndromes, while common regulatory variations are often associated with milder but more frequent forms of the disease. In the most recent Psychiatric Genomics Consortium (PGC) cross-disorder genome-wide association studies (GWAS) meta-analysis [2,3], RNA Binding Fox-1 Homolog 1 (*RBFOX1*) was the second most pleiotropic locus and was found associated with all, but one disorder. This is corroborated by further GWAS on mood and anxiety disorders as well as aggression [4-8]. Beyond these studies on common genetic variation, rare genetic variants in *RBFOX1* such as copy number variants (CNVs) or loss-of-function mutations have been related to autism spectrum disorders (ASD) [9-13]. *RBFOX1* encodes for a splicing factor that regulates the splicing and expression of large networks of genes involved in brain development [31-33]. These data collectively implicate *RBFOX1* as one of the most relevant risk genes for psychopathology, however, neither the specific behavioural domains nor the neural circuits which are involved have yet been identified.

To delineate the genetic contribution of *RBFOX1* to mental disorders, we comprehensively data-mined and synthesized large-scale datasets on common and rare genetic variation in psychiatric disorders and traits (Supplementary Table 1). Genome-wide associations between **single SNPs** in *RBFOX1* and major depressive disorder (MDD; 38 SNPs), risk tolerance (RT; 4 SNPs) and the cross-disorder meta-analysis (CD-MA; 42 SNPs) were found in these studies (Supplementary Table 2). At the **gene level**, *RBFOX1* was found associated with several psychiatric conditions, obtaining again gene-wide significance for MDD ($p = 8.62e-17$), RT ($p = 5.6e-12$), and CD-MA ($p = 1.2e-10$), but this time also for schizophrenia (SCZ; $p = 7.2e-08$), (Figure 1A). Interestingly, genes associated with these disorders were significantly enriched for *RBFOX1* targets (MDD, $p = 0.016$; SCZ, $p = 0.042$; RT, $p = 0.010$; CD-MA, $p = 0.019$) (Supplementary Table 3), as it was previously shown for aggression ($p = 3.4e-05$) [14]. The above evidence therefore highlights *RBFOX1* as a robust, replicated cross-disorder risk gene with pleiotropic effects. Next, we browsed **CNVs spanning *RBFOX1*** reported in patients with psychiatric conditions, identifying CNVs for seven disorders/traits (in total 124 losses and 34 gains). The vast majority of CNVs were found in patients suffering from ASD (112 CNVs), but also in patients suffering from SCZ (24 CNVs) (Figure 1B, Supplementary Figure 1 and Supplementary Table 4), probably due to the

larger number of CNV studies for these disorders. Across all disorders, CNVs were 2.3 times more frequent in cases than in controls, with a notable enrichment in ASD (ratio = 5:1) (Supplementary Table 5). A significant burden of CNVs was observed in two studies on ADHD and ASD (Supplementary Table 5).



**A** Gene-based association
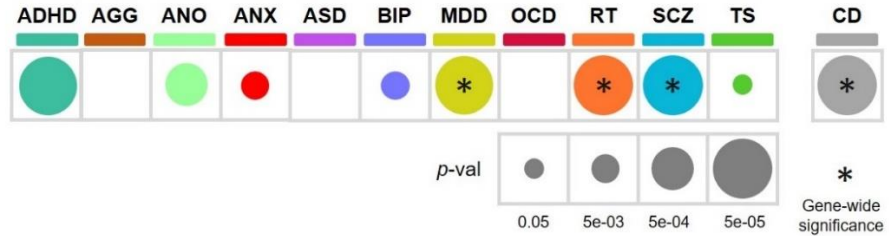
**B** CNVs in ASD and SCZ patients
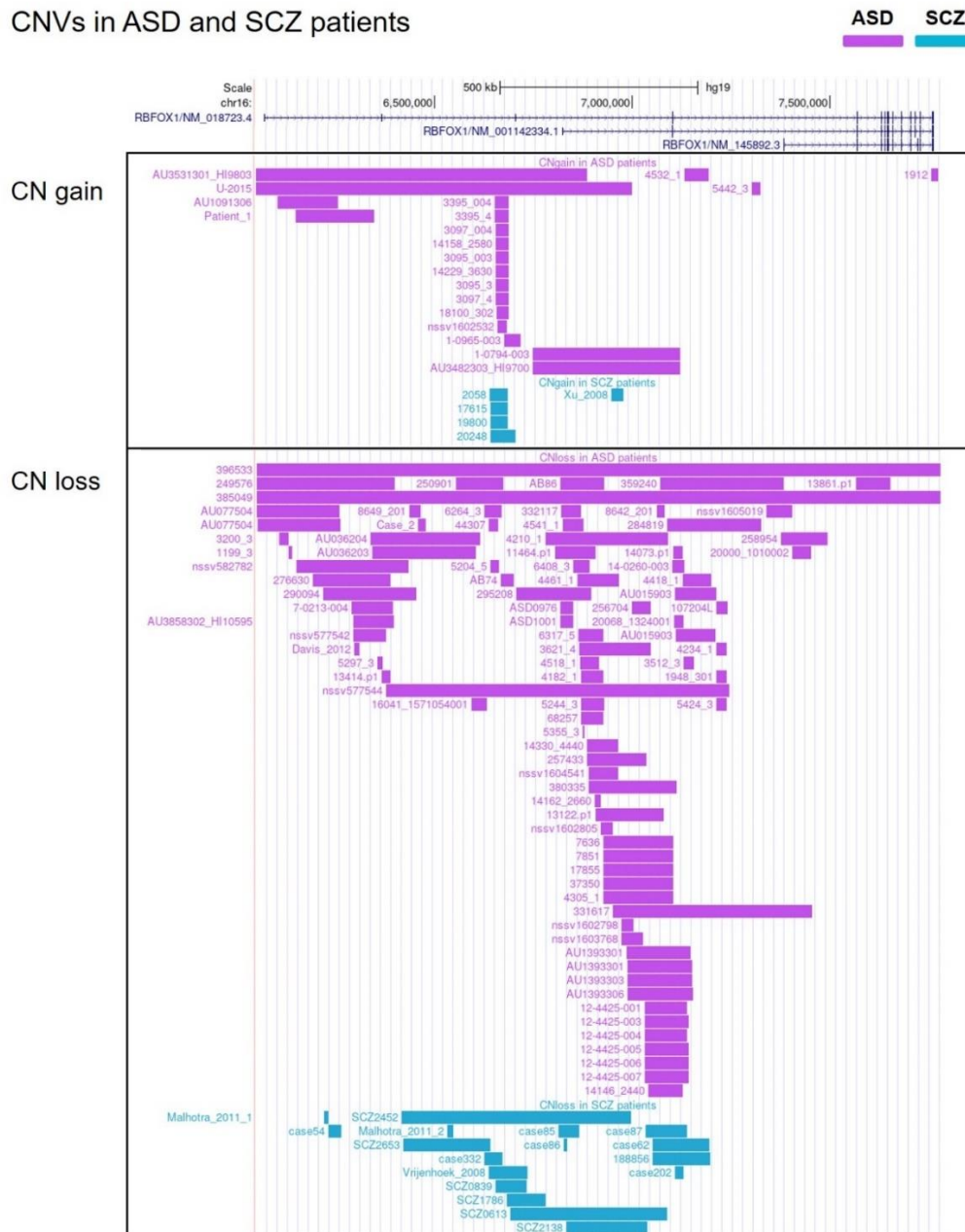
**Figure 1.** Genetic risk variants in *RBFOX1* in different psychiatric conditions and traits. a) Common single-nucleotide variants in *RBFOX1* showed a gene-based association with the majority of disorders and traits. b) Copy number variants (CNVs) identified in ASD and SCZ patients. Top panel, copy number gains identified in ASD and SCZ patients. Bottom panel, CN loss identified in ASD and SCZ patients. Each bar represents a CNV. ADHD, attention-deficit/hyperactivity disorder; AGG, aggression; ANO, anorexia; ANX, anxiety; ASD, autism spectrum disorder; BIP, bipolar disorder; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; RT, risk tolerance behavior; SCZ, schizophrenia; TS, Tourette's syndrome; CD, cross-disorder meta-analysis. p-val, p-value.

These CNVs probably affect RBFOX1 function in distinct ways; while half of them span particular exons or even the whole gene, affecting the coding sequence, many CNVs in introns overlap regions with transcription factors binding activity (intron 2 in gains, and intron 3 in losses) containing putative regulatory elements and potentially altering *RBFOX1* expression (Figure 1B, Supplementary Figures 1, 2 and Supplementary Table 4, 6). In line with this strong evidence that genetically driven variation of *RBFOX1* expression is associated with mental disorders, decreased *RBFOX1* mRNA levels was found in ASD and SCZ patients in cortical regions of *postmortem* brains (Supplementary Table 7). These brain regions converge with those where the expression of *RBFOX1* is highest (Supplementary Figures 3, 4). Taken together, common genetic variation in *RBFOX1* is robustly associated with a variety of mental disorders, while rare genetic variation is especially linked to ASD, where *RBFOX1* expression is found to be decreased in *postmortem* cortical brain tissue.

We then studied the influence of genetic variation in *RBFOX1* on **human neural circuits** to uncover their functional consequences. To do so, we focused on common genetic variants since rare CNVs cannot be examined using neuroimaging or other functional methods with sufficient statistical power. Based on a previous study on aggression, describing an association of *RBFOX1* SNPs with anger, conduct disorder, and aggression [8], we investigated the effect of the most promising SNP (rs6500744; aggression risk allele: C) to circuits underlying emotion processing, fear conditioning, and executive functioning using functional magnetic imaging (**fMRI**), and the *Behavioral Avoidance Task* (**BAT**[15]). Given the role of the anterior cingulate cortex (ACC) in integrating cognition with emotion [16-18], its link with mental disorders [19], and the high level of *RBFOX1* expression in this brain area (Supplementary Figure 3), we assessed the effects of rs6500744 on ACC activation during **executive functioning** and **implicit emotion processing**. Region-of-interest (ROI)-analyses in 313 healthy volunteers showed an increased response of the dorsal ACC for matching fearful as well as angry faces (compared to matching geometric forms) for C-allele carriers compared to T/T carriers (Figure 2B), suggesting increased reactivity to emotional stimuli in the target brain area. No other brain area showed a significant genotype

effect during implicit emotion processing at a stringent whole-brain significance threshold (peak-voxel family-wise error-corrected p < 0.05). In 324 healthy controls, ROI-analyses did not reveal any significant effect of rs6500744 on ACC activation during executive functioning as measured with the Flanker/Go-Nogo task. However, whole-brain analyses revealed that C-allele carriers compared to T/T carriers showed a reduced left dorsolateral prefrontal cortex (DLPFC) response during conditions of cognitive control (contrast [incongruent & nogo] < [congruent & neutral]) (Figure 2A). This reduced DLPFC activation during executive functioning suggests less efficient prefrontal processing during cognitive and impulse control [20,21] that might contribute to both mood disorders and increased impulsivity. Our findings converge with an intermediate DLPFC phenotype related to compromised executive functioning and implicit emotion processing in aggression [22,23]. Altered brain activation during implicit emotion processing and executive functioning as influenced by the effects of *RBFOX1* rs6500744 genotype may therefore underlie the increased risk for psychiatric disorders characterized by increased emotional reactivity (e.g., MDD), impaired impulse control (e.g., ADHD, ASD, risk tolerance), and aggression, all of which are associated with *RBFOX1*.

In an independent dataset, analysed to translate findings from *Rbfox1* knockout mice (see below), we tested whether rs6500744 influences the neural activation in the ACC and amygdala during **fear conditioning** (Figure 2C) in a sample of 47 panic disorder and agoraphobia (PD/AG) patients, again using fMRI. Compared to T/T carriers, C/C carriers revealed a significant activation enhancement in the ACC (Figure 2C and Supplementary Table 10) for simple fear learning, and a significant activation reduction in the dorsal ACC for CS+ after fear extinction (Figure 2C and Supplementary Table 10). ROI analyses with a threshold of p <0.001 within the amygdala did not find any significant genotype differences. Since the dorsal ACC is crucial for fear appraisal [24] and expression [25-27], our findings suggest that rs6500744 C/C genotype carriers display enhanced fear expression after fear conditioning and more fear reduction after extinction training. Notably, these patients also had significantly increased depression (BDI-II) and anxiety (ASI) scores (Supplementary Table 9). In contrast, T-allele carriers fail to demonstrate fear conditioning- and extinction-related changes in neural processing.
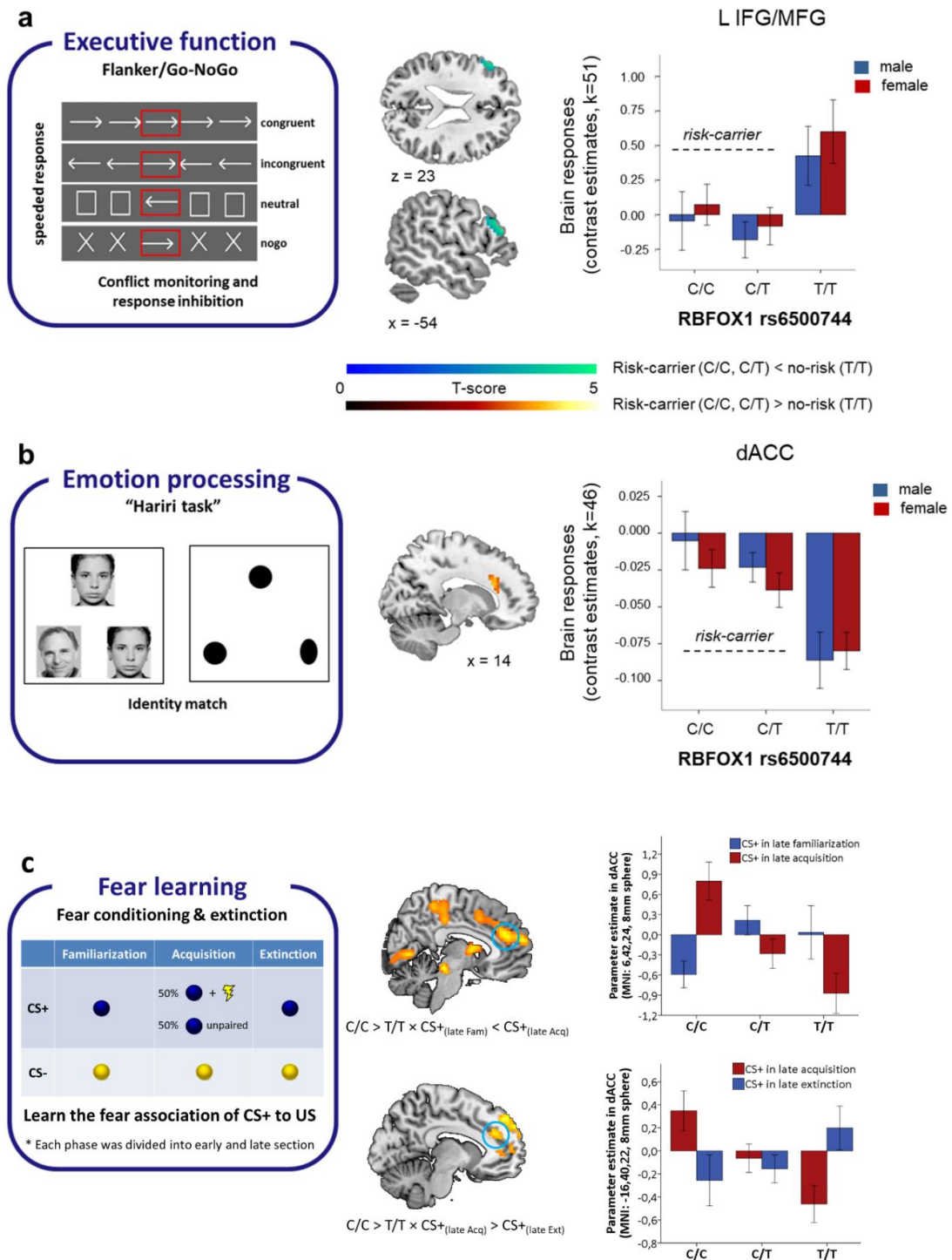
**Figure 2**: Effects of the *rs6500744 RBFOX1* genotype on brain responses during executive functioning, implicit emotion processing in healthy adults, and on fear learning in patients with panic disorder and agoraphobia. **a, left panel**: Schematic overview of the face matching task. Participants had to select either one of the two faces or forms shown at the bottom of the screen that was identical to the target stimulus shown at the top of the screen. **a, right panel**: C-allele carrier (C/C and C/T) showed increased brain responses in the dorsal anterior cingulate cortex (dACC) compared to T/T carrier during matching faces vs. forms (faces>forms; MNI coordinate: x=15, y=23, z=27, peak-voxel family-wise error-corrected [FWE] p = 0.010, *T* = 3.9 within bilateral ACC). **b, left panel**: Schematic overview of the Flanker/Go-Nogo task. Participants had to respond to the direction of the arrow shown in the center (red box for illustration purposes only). **b, right panel**: C-allele carrier (C/C and C/T) showed reduced brain responses in the left

dorso-lateral prefrontal cortex (L dlPFC) compared to T/T carrier during executive functioning (contrast: [incongruent & nogo] > [congruent & neutral]; MNI coordinates: x=-54, y=32, z=21, peak-voxel *pFWE*-corrected=0.039, *T* = 4.55, across the whole-brain). Brain maps were thresholded at *p* < 0.001 uncorrected for display purposes. Error bars indicate ± 1 standard error. **c, left panel:** Schematic overview of the fear conditioning and extinction task. During the acquisition phase, 50% of CS+ was paired pseudo-randomly with the US and 50% were not. Only those trials in which no US was delivered were analyzed during acquisition to avoid overlap with neuronal activation directly related to the presentation of the US. **c, right panel:** Using ROI analysis within the ACC, homozygote risk allele carrier (C/C) compared to T/T carrier revealed increased activation in the dorsal ACC for CS+ after fear acquisition (CS+ in the late acquisition > CS+ in the late familiarization; cluster size = 160; $p_{FWE-corrected}$ = 0.009), and activation reduction for CS+ after fear extinction (CS+ in the late acquisition > CS+ in the late extinction; cluster size = 64; $p_{FWE-corrected}$ = 0.04). Brain maps were thresholded at *p* < 0.001 uncorrected for display purposes. Error bars indicate ± 1 standard error.

To further investigate the effect of the rs6500744 genotype on **fear behaviour**, we examined its effect on avoidance during the BAT, where a behavioural and autonomous response to a fear-inducing situation is measured, in 333 PD/AG patients. The rs6500744 C-allele was significantly and dose-dependently associated with the frequency of avoidance behavior (T/T: 13/58 patients, 22.4%; T/C: 39/156 patients, 25.0%; C/C: 44/75 patients, 37.0%; linear trend: p = 0.022). This result was concurrent with observed differences between genotypes according to everyday life avoidance behaviour, assessed by clinical expert ratings (Clinical Global Index): again, avoidance increased linearly with the number of C-alleles (T/T: m=4.31, SD=1.29; T/C: m=4.59, SD=1.10; C/C: m=4.65, SD=1.03; linear trend: p = 0.04). In those 106 BAT non-avoiding patients with at least moderate fear during the task, the heart rate during both anticipation and exposure phase significantly increased relative to recovery phase with an increasing number of C-alleles (linear trend BAT phase × genotype: p = 0.031, Supplementary table 12) indicating increasing autonomic threat processing. Importantly, T-allele homozygotes did not show *any* heart rate modulation during the BAT. Together with the fMRI data, this suggests that rs6500744 C-allele carriers show more avoidance behaviour due to better fear learning and stimulus discrimination.

The molecular consequences of the SNPs and CNVs in *RBFOX1* are yet unknown, although a numerically, but not statistically significant decrease of RBFOX1 protein in rs6500744 T-allele carriers (Supplementary Figure 5) was found in human *postmortem* frontal cortical tissue. However, given the decreased expression of *RBFOX1* observed in *postmortem* studies of ASD and SCZ patients, and the over-abundance of *RBFOX1* CNV-deletions in mental disorders, we reasoned that loss of RBFOX1 function might underly the observed associations and generated a neuron-specific *Rbfox1* knockout (KO) mouse line to determine the behavioural consequences of decreased *Rbfox1* expression. In line with the human CNV studies, we observed **pronounced**

**hyperactivity** in the KO mice, which was observed in the open field, light dark box (Figure 3A-C) and marble burying tests (Supplementary Figure 6A). Interestingly, this was coupled with thigmotaxis and **stereotypic**-like behaviour, as the KO spend double the time adjacent to the maze walls (Figure 3A). These behaviours confound the typical measures of anxiety in these tests, as KOs spent three times longer investigating a novel object placed in the centre of the OF (Figure 3B). While KO mice showed a deficit in the acoustic startle response (Figure 3D), sensorimotor gating was not impaired.

Next, we moved to more cognitively demanding tests, such as the ones performed in the human cohorts. Here we found evidence for gene-dose dissociation, as both heterozygous (HET) as well as KO mice had impairments in fear acquisition and extinction in the auditory cued **fear conditioning** test (Figure 3F). In contrast, compared to KO, HET mice could acquire fear conditioning similar to littermate controls, but they were unable to retain fear memory. This deficit was specific for cued fear learning, as neither associative reward learning in a touchscreen pairwise visual discrimination task (Figure 3G; Supplementary Figure 6B) nor spatial learning (spontaneous alternations in the Y-maze) were impaired (Figure 3E).

Finally, given the strong genetic association of *RBFOX1* with ASD and aggression, we assessed **social interaction** as well as male-male **aggression**. Supporting genetic evidence, we observed significantly less social interest in the KO mice (Figure 3H) which also manifested in a lack of aggressive behaviour (Figure 3I). Thus, neuron-specific *Rbfox1* depletion in mice leads to hallmark features of ASD, namely, repetitive-stereotyped hyperactive locomotor behaviour, abnormalities in the fear circuitry, and impaired social interactions [28]. Such pronounced effects of neuron-specific loss of *Rbfox1* might thus also occur in human carriers of rare loss-of-function variants with high penetrance, underscoring the relevance of this gene for neurodevelopment.

Collectively, the evidence from genetic studies accrued here suggest that common genetic variation of *RBFOX1* goes along with a wide spectrum of psychiatric phenotypes, while rare CNVs in this gene contribute especially to ASD, although this might be biased by the low number of studies investigating CNVs in other psychiatric disorders. The molecular-cellular effects of common genetic variation in *RBFOX1* are however yet elusive and likely include the affectation of regulated gene networks. This may be operative only in certain cell types or developmental stages, as the some of the major roles of RBFOX1 occur during early brain maturation [29], where it orchestrates downstream genetic networks implicated in neuronal development [30] via direct regulation of post-transcriptional programs . These gene networks are markedly inter-connected and enriched for genes relevant for cortical development and autism [31] as well as MDD and SCZ (Supplementary Table 3) susceptibility. On target transcripts, *RBFOX1* regulates

alternative splicing of tissue-specific exons [32] by binding to mRNA GCAUG motifs in the nucleus and affecting mRNA stability in the soma, and thus has different roles in those intracellular compartments. Importantly, it promotes interneuron-specific connectivity in the developing neocortex [33] by regulating cell-type-specific splicing (parvalbumin [PV] vs. somatostatin [SST] interneurons). Loss of *RBFOX1* in inhibitory interneurons causes significantly reduced synaptic transmission [34], by affecting membrane excitation and neurotransmission [35], resulting in reduced inhibition of the postsynaptic neuron and leading to excitatory/inhibitory (E/I) imbalance, a key feature of ASD. As PV+ interneurons are regulators of E/I balance [36], this might link dysregulation of *RBFOX1* to E/I dysbalance to ASD susceptibility.
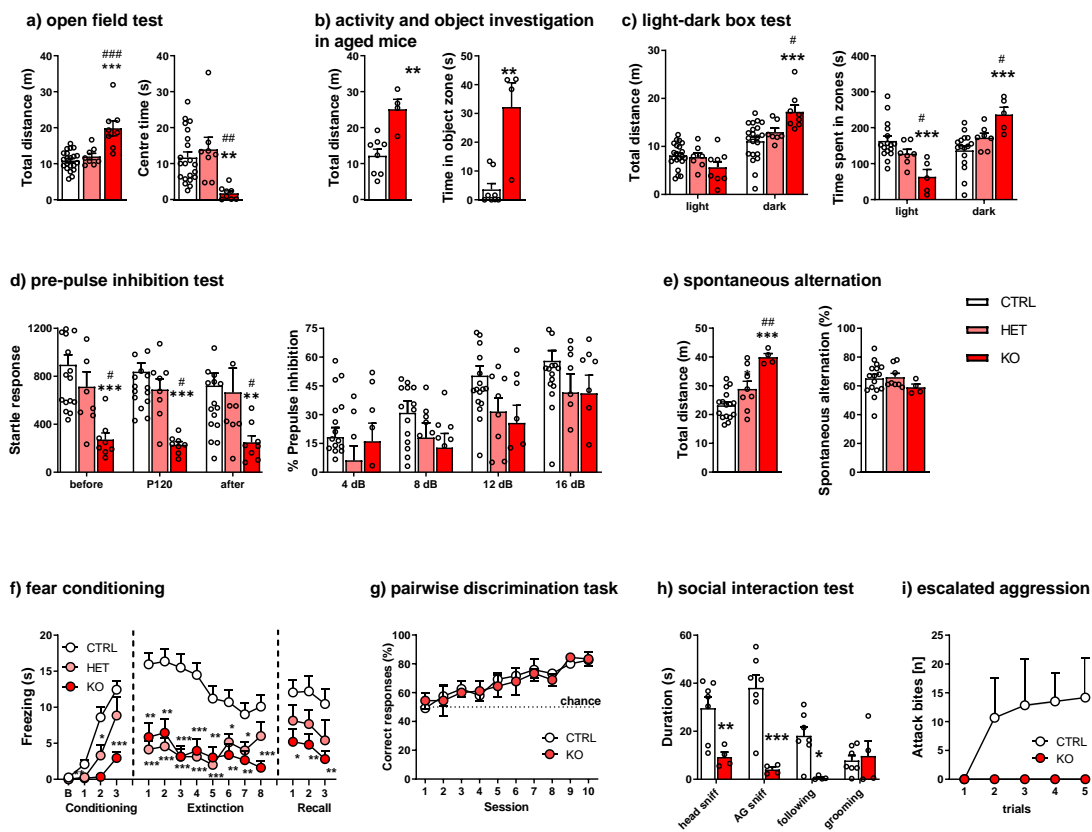


**Figure 3.** Effects of neuronal-specific *Rbfox1* deletion on behavioural measures in male mice. **a,** open field test: *Rbfox1*-KO mice displayed hyperactivity and thigmotaxis in the open field test; **b,** open field test and novel object exploration in aged mice: KO mice spent longer investigating a novel object placed into the open field; **c,** light-dark box test: KO mice again were hyperactive and spent less time in the light zone; **d,** pre-pulse inhibition test: KO mice had markedly reduced startle amplitude; **e,** spontaneous alternation: the number of spontaneous alternations was not changed in KO although the distance travelled during the test was significantly higher; **f,** auditory fear conditioning: fear acquisition and extinction was impaired in the KO mice, and HET mice displayed impaired fear retention; **g,** touchscreen visual pairwise discrimination task: acquisition of the task was similar in CTRL and KO; **h,** social interaction: KO spent significantly less time investigating unfamiliar stimulus mice; **i,** escalated aggression paradigm: while aggressive behaviour increased during repeated sessions in CTRL, KO remained non-aggressive

throughout testing. Data is presented as means ± S.E.M. * p < 0.05; ** p < 0.01; *** p < 0.001 vs CTRL; # p < 0.05; ## p < 0.01; ### p < 0.001 vs HET.

With respect to common genetic variation, *RBFOX1* is associated with all disorders combined, SCZ, MDD, and RT. Our neuroimaging data however argues for an effect of *RBFOX1* genetic variation on the networks controlling fear learning, executive functioning, and emotional processing. rs6500744 risk genotype carriers display higher reactivity to emotional stimuli and reduced DLPFC activation during cognitive control, which are both linked to these mental disorders. Increased aggression found in C-allele carriers [8] is thus likely to be interpreted as reactive-impulsive, but not proactive, aggression. It must be considered that genetic variants in *RBFOX1* with small effect sizes in a polygenic scenario interact with many other variants to increase the risk towards mental disorders in a quasi-stochastic manner, probably explaining the broad psychopathological phenotype. In contrast, more penetrant CNVs with presumably stronger molecular effects may result in a more specific chronic neurodevelopmental behavioural syndrome.

While we cannot yet finally determine the functional consequences of *RBFOX1* genetic variation in humans, combined data from human and rodent experiments would rather postulate an increase in expression in MDD, anxiety and (reactive) aggression. Up- and downregulation of *RBFOX1* however might have different effects on the regulated gene networks [37], and human *postmortem* data argues for reduced *RBFOX1* expression at least in ASD and SCZ. In line with this hypothesis, the remarkable behavioural phenotype of neuron-specific *Rbfox1* knockout mice suggests that loss-of-function of *RBFOX1* causes a behavioural syndrome characterized by hyperactivity, stereotypies and specific cognitive and social impairments typical for ASD. We propose that the clinical phenotype in human *RBFOX1* CNV carriers, extending beyond "pure" ASD, is additionally shaped by genetic background and environmental factors. Given the excellent construct and face validity of our *Rbfox1* KO mouse, we consider it as an excellent animal model for ASD with an unprecedent robust behavioural phenotype.

Such differential consequences of common and rare genetic variation may be a general principle in psychiatric genetics, where common variation might underly more unspecific vulnerability, while rare, highly penetrant variation causes more specific phenotypes. In either case, it becomes clear that current diagnostic boundaries do not adequately reflect corresponding biological disease types. Given that approaches to modify *RBFOX1* expression are already at hand, which might be used in the sense of personalized mental health, this calls for mechanistic rather than atheoretical, operationalized definitions of mental disorders.

## MATERIAL AND METHODS

### Common and rare genetic risk variants in *RBFOX1* in psychiatric phenotypes

The contribution of common variants (MAF > 0.01) in the *RBFOX1* gene to psychiatric disorders or related behavioural traits was assessed through SNP-based and gene-based association studies using GWAS summary statistics from previous studies (Supplementary Table 1). In total, eleven psychiatric conditions or traits were investigated: attention deficit-hyperactivity disorder (ADHD), aggression (AGG), anorexia (ANO), anxiety (ANX), autism spectrum disorder (ASD), bipolar disorder (BIP), major depressive disorder (MDD), obsessive-compulsive disorder (OCD), risk tolerance (RT), schizophrenia (SCZ) and Tourette's syndrome (TS), and the cross-disorder meta-analysis of eight of them (CD-MA).

For SNP-based analysis of *RBFOX1* (NM_018723: chr16:6,069,132-7,763,340, GRCh37/hg19 UCSC RefSeq), we included a flanking region of 10 kb at 5' and 5 kb at 3' of the gene and retrieved information of all suggestive associated SNPs (p < 1e-05) from each summary statistics dataset. Gene-based association studies were performed on MAGMA v1.06 using the SNP-wise mean model without window around the gene, and the 1000 Genomes Project Phase 3 (European data only) as a reference panel. The enrichment of *RBFOX1* target genes was assessed by a hypergeometric test.

For copy number variants (CNVs) in *RBFOX1*, we collected publicly available data from the above disorders or traits (in patients and in controls when reported), either in published papers (until April 2020) or databases (DECIPHER, https://decipher.sanger.ac.uk; ClinVar, https://www.ncbi.nlm.nih.gov/clinvar; ISCA, http://dbsearch.clinicalgenome.org/search/). To inspect overlap between CNVs identified in patients and putative cis-regulatory elements we used epigenetic data from ENCODE (https://www.encodeproject.org/) from seven neural tissues and brain-related Hi-C data from 3DIV to identify interactions with the first distal promoter. We performed burden analysis for *RBFOX1* CNVs in 18 out of 34 studies where information in controls was available using PLINK v.1.07 considering CN loss and CN gains separately as well as both together.

### Expression of *RBFOX1* in brain samples of ASD and SCZ patients

Alterations in the expression of *RBFOX1* in the brain was assessed using transcriptomic data from post-mortem brain samples of ASD and SCZ patients, compared to controls, using publicly available human datasets, either in GEO (http://www.ncbi.nlm.nih.gov/geo) or published

articles (Supplementary Table 7). *RBFOX1* expression was explored in different brain areas, including hippocampus, cerebellum or cortex, depending on the dataset.

**RBFOX1 rs6500744 effect on RBFOX1 protein expression in *post-mortem* tissue samples**

Human *post-mortem* tissue samples were obtained from the NeuroBioBank of the National Institute of Health (NIH NBB; Request ID #613) in accordance with institutional and ethical guidelines. Age-matched samples from the dorsolateral prefrontal cortex (DLPFC) of psychiatrically healthy male and female subjects from five biorepositories were genotyped for the *RBFOX1* rs6500744 using Kompetitive Allele-Specific PCR (KASP) (Supplementary Tables 14 and 15). Equal amounts of protein from the tissue lysates were subjected to Western blotting to detect RBFOX1 expression.

**RBFOX1 rs6500744 in functional MRI**

*Genotyping*

DNA was extracted from whole blood according to standard procedures for all participants. Then, genome-wide SNP (single nucleotide polymorphism) genotyping was performed using a standard GWAS chip (PsychChip XY; Illumina Human610-Quad BeadChip [Illumina, Inc., San Diego]). Based on this genome-wide chip, genotype information for the rs6500744 *RBFOX1* SNP was retrieved for each individual using plink (http://zzz.bwh.harvard.edu/plink/). The observed genotype distribution of *rs6500744* did not deviate from the Hardy-Weinberg equilibrium (Flanker/Go-Nogo: $p = 0.826$ [C/C carrier: $n = 71$, C/T carrier: $n = 158$, T/T carrier: $n = 95$]; Face matching: $p = 0.821$ [C/C carrier: $n = 70$, C/T carrier: $n = 154$, T/T carrier: $n = 89$]; computed based on the CRAN R-package, https://cran.r-project.org/web/packages /Hardy-Weinberg /index.html;). Analogous to previous imaging genetic studies on common genetic risk variants for psychiatric disorders (e.g., [38,39-41]), we compared brain activation for risk-allele carrier (C/C and C/T) to no-risk allele carrier (T/T carrier) of rs6500744 (as proposed by Fernandez-Castillo et al 2020).

*Flanker/Go-NoGo and Face matching tasks*

*Sample*: We included the data of 324 (Flanker/Go-Nogo task) and 313 (Face matching task) healthy adults of European ancestry who have been recruited as healthy controls within the framework of a multi-site imaging genetics study assessing the intermediate phenotypes of

psychiatric disorders such as depression, schizophrenia, and bipolar disorder (for previous work, see [39,42,43-46]). Data collection was carried out at the Central Institute of Mental Health in Mannheim, at the Medical Faculty of the University of Bonn, and the Charité University Medicine in Berlin. All participants provided a whole-blood sample for DNA extraction and underwent a well-established implicit emotion processing paradigm (face matching task [47]) and a Flanker/Go-Nogo task [48] during fMRI. All participants provided written informed consent for study protocols approved by the Ethics committees of the Medical Faculty of Mannheim at the Ruprecht-Karls-University in Heidelberg, the Medical Faculty of the University of Bonn, and the Charité University Medicine in Berlin. General exclusion criteria for the healthy controls were a lifetime history of significant general medical, psychiatric, or neurological disorders, a family history of psychiatric disorders, current or past psychotropic pharmacological treatment, drug or alcohol use as well as head trauma (compare also, [39,46]).

*Task procedures:* In each trial of the Flanker/Go-NoGo task, participants saw an array of five stimuli that included a central target arrow pointing left or right, flanked by two stimuli (arrows, boxes, or Xs, see **Figure 2a, left panel**) on either side. Participants were instructed to press a button corresponding to the direction of the central target arrow as fast and accurately as possible. In the "conflict monitoring" condition, the flanking arrows pointed either in the same direction (congruence, n=41 trials) or the opposite direction of the central arrow (incongruence, n=40 trials); the incongruence condition has been shown to slow down response times and indicates attentional capability to resolve conflict. In the "neutral" condition (n=31 trials), the central arrow was flanked by boxes, measuring response execution without any conflict. In the "nogo" condition (n=33 trials), the central arrow was flanked by "Xs" which indicated that the participants had to withhold their response. The nogo condition is an established method to measure response inhibition. Each stimulus combination was presented for 800 milliseconds (ms) followed by a variable inter-trial interval of 2200–5200ms in which a fixation cross was shown in the centre of the screen. Task performance was examined by the accuracy for each condition (% correct) and the reaction time (RT in ms) for congruent, incongruent and neutral conditions. All participants included in the current data analysis had an accuracy >60% for each condition. For each participant, brain activation was estimated for each task condition by computing a general linear model using SPM8 including regressors for the congruent, incongruent, nogo and neutral conditions as well as six realignment parameters (3 translation, 3 rotation). We convolved a stick function representing the onsets of trials with the SPM8 canonical hemodynamic response function.

For the Face-matching task, participants in each trial viewed either a trio of angry or fearful faces or neutral geometric forms (see **Figure 2b, left panel**). They were instructed to select one of the two faces or forms presented at the bottom of the screen that was identical to the target stimulus presented at the top of the screen. In total, there were eight blocks with six images that were presented sequentially for 5 seconds, either three faces of each target affect (angry or fearful) and gender or six neutral forms. Task performance was assessed as a percentage of correctly answered trials for the face and form matching condition. For each participant, brain activation was estimated for each task condition by computing a general linear model using SPM8 including regressors for the face matching and the form matching blocks as well as six realignment parameters (3 translation, 3 rotation). We convolved a boxcar function representing the duration of blocks with the SPM8 canonical hemodynamic response function.

*fMRI*

Functional MRI data were acquired on three comparable 3T TrioTim MRI scanners (Siemens, Erlangen, Germany) in Mannheim, Bonn and Berlin using a gradient-recalled echo-planar imaging sequence (GRE-EPI) with the following MR parameters: 28 axial slices per volume, 4 mm slice thickness, 1 mm gap, TR = 2000 ms, TE = 30 ms, field of view (FOV) = 192 mm, flip angle = 80°, acquired in descending order. We acquired 135 volumes for the face matching task and 306 volumes for the Flanker/Go-Nogo task. Additionally, high-resolution T1 structural data were acquired using a 3D magnetization-prepared rapid gradient-echo (MP-RAGE) sequence with the following sequence parameters: 176 sagittal slices, 1 mm slice thickness, TR = 1570 ms, TE = 2.75 ms, TI = 800 ms, FOV = 256 mm, flip angle = 15°. Preprocessing and estimation of functional task-dependent brain activation at the subject level were carried out using the MATLAB-based statistical parametric mapping software (version SPM8, Wellcome Trust Centre for Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk/spm/). Functional images were preprocessed for each participant. fMRI data were slice time corrected, realigned to the first image of the time series, spatially normalized to the Montreal Neurological Institute (MNI) template, resampled to 3 mm isotropic voxels, and smoothed with a 9 mm full-width at half-maximum (FWHM) Gaussian filter. Second-level analyses testing for genotype effects across participants were carried out using SPM12.

Individual brain activation maps were subjected to separate 3 (*rs6500744* genotype: C-carrier, T/T carrier) x 2 (sex: men, women) full-factorial models including age and imaging site as regressors of no interest using SPM12 to test for the effects of *rs6500744* on brain responses during response inhibition ("nogo > neutral" contrast), conflict monitoring ("incongruent >

congruent" contrast), overall executive functioning (combined contrast: [nogo & incongruent] > [neutral & congruent]), and implicit emotion processing ("faces > forms") for second level fMRI analyses. Sex was included as a between-subject factor into the full-factorial model to identify potential genotype x sex interactions in imaging space due to significant main effects of sex and sex by genotype interactions on behavioural performance for both tasks and previously reported sex by genotype effects for comparable intermediate phenotypes for genetic variation of the *MAOA* gene [48]. Given that altered ACC functioning during executive functioning measured with the Flanker/Go-Nogo task and during implicit emotion processing measured with the face matching task has previously been associated with different psychiatric risk genotypes (e.g., *MAOA*, *5-HTTLPR*, *BDNF Val[66]MET*) [38,39,48], we tested genotype effects in an a *priori* defined standard anatomical mask of the ACC derived from the Automated Anatomical Labeling (AAL) atlas [49]. The ACC AAL mask covers the dorsal cognitive and rostral-ventral affective divisions of the ACC including the areas 24 a-c, 25, 32, 33, and parts of the areas 24a'-c' and 32' (located below z=31) as defined by Bush and colleagues [50], or the areas adACC (below z=31), pgACC and sgACC according to Etkin and colleagues [24] (total mask volume: 22.032 mm$^3$, maximum extensions in MNI space: x = -16 to 19, y = -4 to 55, z = -10 to 30). The significance level was set to $P < 0.05$ family-wise error (FWE) corrected for multiple comparisons across all voxels within the ACC AAL mask. Outside this pre-hypothesized ROI, findings were considered significant if they passed a significance threshold of $P < 0.05$ FWE corrected for multiple comparisons across the whole brain.

*Fear conditioning and extinction*

The MAC multicenter psychotherapy study recruited in total 369 patients of Caucasian origin met DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) criteria for PD/AG, as assessed by the Composite International Diagnostic Interview. An overview of the number of participants for every substudy (genetic data, psychophysiological assessment, and functional magnetic resonance imaging (fMRI)) is given in Supplementary Table 9. The study with all of its subprojects was approved by the respective local ethical committees and informed consent was obtained.

A total of 47 patients with valid fMRI data sets were sequenced regarding their *RBFOX1* SNP *rs6500744*. The distribution of CC, CT and TT genotype is 15, 21 and 11, respectively. A differential fear conditioning procedure was applied to probe neural correlates of fear acquisition. Differential fear conditioning task during fMRI scanning consisted of three phases [Familiarization (F) with 16 trials; acquisition (A) with 32 trials and extinction (E) with 16 trials of

each conditioned stimulus (CS)+ and CS– (coloured geometrical forms); presentation time: 2,000 ms with a variable inter-trial interval (ITI) of 4.785–7.250 s]. Unconditioned stimulus (US), an aversive tone (100 ms white noise between 70 to 105 dB), was firstly familiarized (16 trials of isolated presentation), then pseudo-randomly paired with one of the CSs (counterbalanced between subjects; partial reinforcement rate of 50 %) during acquisition, resulting in equal proportions of CS+$_{paired}$ and CS+$_{unpaired}$ trials. Detailed information on the fMRI task, data acquisition, analysis pathway and data quality control is provided elsewhere [51,52] and in the online resource. For all analyses, voxels with a significance level of $p < 0.005$ uncorrected belonging to clusters with at least 142 voxels are reported.

## *Behavioural avoidance task*

The BAT (behavioural avoidance task) assessment was part of two study waves of the German national research network PANIC-NET. Genotyping for *rs6500744* was available for a total of 333 participating patients (n(CC)=119; n(TC)=156; n(TT)=58) with a primary DSM-IV-TR diagnosis of panic disorder with agoraphobia with at least moderate disorder severity (238 females; mean age: m=35.50 years, SD=10.72; no significant differences between genotype groups). Criteria of inclusion and exclusion and patient recruitment procedure are described in detail elsewhere [53]. Diagnosis of PD/AG was established by a standardized computer-administered face-to-face Computer Assisted Personal Interview-World Health Organization-Composite International Diagnostic Interview (CAPI-WHO-CIDI) by trained and certified interviewers. All patients were free from psychotropic medication. Patients gave written informed consent after receiving a detailed description of the study. The study was approved by the Ethics Committee of the Medical Faculty of the Technical University of Dresden, which was valid for all participating centres. The highly standardized BAT procedure is described in detail elsewhere [15]. Briefly, after an anticipatory phase (sitting in front of the open test chamber) patients were asked to stay in a small, closed and dark test chamber for a maximum of 10 minutes (unknown for the patients). During BAT exposure the patients could refuse or end the test prematurely (passive and active avoidance behaviour, respectively) at any time. After BAT exposure a recovery phase followed (again sitting in front of the open test chamber). Reported fear was assessed immediately after each period (Likert scale ranging from 1 to 10). Heart rate was calculated from a continuously recorded electrocardiogram. Due to technical failures, heart rate was available only in a subsample of patients. To test for a significant association between *rs6500744* genotype and BAT avoidance behaviour a chi-square test was conducted. The genotype effect on heart rate response was tested applying a mixed model of variance including genotype as a between-subject factor and BAT phase as within-subject factor.

**Animals**

Male *Rbfox1*<sup>fl/fl</sup> (*Rbfox1*<sup>tm1.1Dblk/J</sup>; JAX strain 014089) and *Synapsin1*-Cre (B6.Cg-Tg(Syn1-cre)671Jxm/J; JAX strain 003966) mice were maintained on a C57Bl/6J background and housed in groups of 2-5 in standard individually ventilated cages on a 12 h light/dark cycle (lights on at 7:00) under controlled ambient conditions (21±1°C, 55±5% humidity). Food and water were available *ad libitum* unless specified otherwise. To generate mice with neuronal-specific deletion of *Rbfox1* (*Rbfox1*-KO), *Rbfox1*<sup>fl/fl</sup> mice were crossed to mice carrying Cre-recombinase under the direction of the rat *Synapsin* I promoter (*Synapsin1*-Cre). The resulting heterozygous *Rbfox1*<sup>fl/+</sup>/*Synapsin1*-Cre<sup>+/-</sup> (HET) female mice were crossed to *Rbfox1*<sup>fl/fl</sup> (CTRL) males to produce homozygous *Rbfox1*<sup>fl/fl</sup>/*Synapsin1*-Cre<sup>+/-</sup> (KO) offspring. *Rbfox1*<sup>fl/fl</sup>/ and *RBFOX1*<sup>fl/+</sup>/ Synapsin-Cre<sup>-/-</sup> mice were used as controls. Male C57Bl/6J mice were used as social stimuli in the social interaction test and intruders in the aggression testing paradigm. All breeding and experimental procedures were conducted in accordance with the Directive of the European Communities Council of 24 November 1986 (86/609/EEC) and German animal welfare laws (TierSchG and TSchV) and were approved by the Darmstadt regional council (approval ID: FK/1126).

**Behavioural experiments.**

For habituation purposes, mice were transported to the behavioural testing room at least 45 min before testing. Experiments were performed between 9:00 – 14:00. Behavioural apparatuses were cleaned before testing and between animals using Aerodesin 2000 (Lysoform Dr Hans Rosemann GmbH, Berlin, Germany).

*Open field test (OF) and novel object investigation.* The OF apparatus (Stoelting Europe, Dublin, Ireland) consisted of a 40 cm x 40 cm grey arena surrounded by black perspex walls (height: 35 cm) and a USB camera (The Imaging Source Europe, Bremen, Germany) was fixed on a metal arm above it. Each mouse was placed into the centre of the OF and allowed to explore freely for 3 min. The arena was virtually divided into a 15 cm x 15 cm centre area and a 10 cm wide periphery. Distance travelled and time spent in the centre was automatically quantified using ANY-maze automated tracking software (Stoelting Europe, Dublin, Ireland). Each mouse was initially tested in the OF at the age of 10-12 weeks. At the age of 8-9 months, the same animals were re-assessed in the OF. After a 5-minute OF exposure, a small glass jar was placed into the centre of the OF, and visits to the object zone (ca 2 cm radius around the glass jar) were quantified using ANY-maze.

*Light-dark box (LDB).* The LDB apparatus (Stoelting Europe, Dublin, Ireland) consisted of an enclosure (W: 40 cm x L: 40 cm x H: 35 cm) divided equally into two 20 x 40 cm compartments: a brightly lit clear acrylic glass compartment (illumination level: ~400 lux) and a dark IR-transparent black perspex compartment (illumination level: ~3 lux by an infrared light source) which were connected by an opening in the centre wall. Mice were individually placed into the light area and allowed to freely explore for 5 min. Distance travelled and time spent in each area, transitions between the compartments, and latencies to exit from and re-enter the light area were recorded using an IR-sensitive USB camera (The Imaging Source Europe, Bremen, Germany) and automatically tracked using ANY-maze software (Stoelting Europe, Dublin, Ireland).

*Touchscreen pairwise visual discrimination and reversal learning task.* Pairwise visual discrimination task and reversal learning were conducted in four operant chambers (Campden Instruments Ltd, Loughborough, UK) fitted with touchscreens that displayed the visual stimuli controlled by ABET II Software (Lafayette, IN, USA). The chambers were also equipped with liquid reward (strawberry milk, Müllermilch Erdbeer, Müller, Aretsried, Germany) dispensers, light and sound generators, and USB cameras, and were housed in sound-attenuated ventilated cubicles. Before the start of the experiments, mice were placed on mild food restriction for one week to obtain a weight reduction of 5-10 % of their free-feeding weight. Pre-training consisted of several steps to gradually shape the required behaviours to perform the task (habituation to the test chambers, screen-touching, receiving the reward, initiating trials).  Visual discrimination was performed as described previously [54]. Briefly, the task consisted of pairwise discrimination of a rewarded (S+, "fan") and unrewarded (S-, "marbles") black and white images. The location of the S+ presentation was pseudo randomised. Touching the S+ triggered reward delivery while responding to the S- started a 5-s timeout with the house light on and no reward delivery. Sessions lasted until the animal completed 30 trials or after 60 min, whichever occurred earlier. The criterion for task acquisition was ≥ 80% accuracy (correct responses at S+) during the trial for two consecutive sessions.

*Spontaneous alternation in the Y-maze.* Spontaneous alternation was assessed in an apparatus (Stoelting Europe, Dublin, Ireland) which consisted of three identical arms (35 cm x 5 cm with 10 cm high walls) mounted in the shape of 'Y'. Each mouse was placed into one of the arms and allowed to explore each of the three arms freely for 5 min. A spontaneous alternation occurred when a mouse visited different arms on each of the last three arm entries. The total number of arm entries and spontaneous alternations as well as total distance travelled in the Y-maze were

recorded via a USB camera and quantified using ANY-maze. The percentage of spontaneous alternations of all arm entries was calculated according to the following formula:

$$Spontaneous\ alternation\ (\%) = 100 \times \frac{spontaneous\ alternations}{total\ arm\ entries - 2}$$

*Prepulse inhibition (PPI) of the acoustic startle response (ASR).* PPI of the ASR was measured in the SR-LAB™ startle response system (San Diego Instruments, Inc., USA). Briefly, after 5 min acclimation to the background noise (65 dB white noise), mice were exposed to six startle pulse trials (120 dB broadband noise for 40 ms, 10 s inter-trial interval [ITI]). Then, mice were presented with 10 x no-stimulus, 10 x startle pulse, 10 x each prepulse (4, 8, 12, 16 dB above background = 69, 73, 77, 81 dB for 20 ms) followed after 80 ms by a startle pulse, 10 x prepulse only (81 dB) in pseudorandomized order with a variable ITI (20-30 s). The test session ended with six startle pulse trials separated by 10 s ITIs. The overall duration of this protocol was ca 35 min. The magnitude of the ASR (whole body reflex) to pulse only trials was averaged for each mouse and defined as startle amplitude. Percentage of PPI was calculated as described in Esen-Sehir et al. (2019) using the following formula:

$$PPI(\%) = 100 \times \frac{startle\ amplitude_{startle\ trials} - startle\ amplitude_{prepulse+startle\ trials}}{startle\ amplitude_{startle\ trials}}$$

*Cued fear conditioning and extinction.* Fear conditioning was conducted in the Ugo Basile fear conditioning system (Ugo Basile S.R.L, Gemonio, Italy) over three consecutive days. On day 1, for the acquisition of conditioned fear, mice were placed into a chamber with transparent walls and an electrified grid floor inside a sound-attenuated ventilated cubicle and allowed to habituate for 3 min. Fear conditioning was conducted with three pairings of a 30-s 80-dB 1000-Hz sound (conditioned stimulus, CS) terminating with a 2-s 0.4-mA scrambling footshock (inter-trial interval: 2 min). 24 hours later, on day 2, the conditioning chamber was transformed with black and white cardboard panels on the walls and a grey opaque perspex floor panel covering the electric grid. For the fear extinction test, mice were individually placed into the chamber and allowed to habituate for 3 min, followed by 16 30-s CS presentations with 5-s intervals. Another 24 hours later, fear extinction recall was assessed. The conditioning chamber was set up identically to that of the day before. Mice were individually placed into the chamber, allowed to habituate for 3 min, and exposed to three 30-s CS with 5-s intervals. As an index of conditioned fear, freezing behaviour (defined as inhibition of all movement except breathing) was recorded via a USB camera and automatically quantified using ANY-maze freezing detection module (threshold: 1000 ms).

*Marble burying.* 18 glass marbles (with a diameter of 15 mm) were evenly placed onto the clean bedding material in a clean cage. Each mouse was placed into the cage with marbles and allowed to explore for 30 min. Every 5 minutes, the number of buried marbles (more than 2/3 covered by bedding) was counted.

*Resident-intruder test.* Aggressive behaviour was assessed using the escalated aggression paradigm of repeated daily resident-intruder tests. To increase the territoriality of the residents (the experimental animals), mice were pair-housed with an 8-week-old C57Bl/6J female mouse for one week before the aggression testing. The cages were not cleaned until after the aggression testing was finished. Before the daily testing, the female and any nesting material were removed from the cage. A juvenile intruder (5-week-old male C57Bl/6J) mouse was placed into the home cage of the resident and aggressive behaviour (attack latency and attack frequency) of the resident towards the intruder was manually scored for 5 min after the first attack. If no attack occurred within the first 5 min, the test was ended. Resident-intruder tests were conducted for five consecutive days.

*Social interaction test.* The social interaction test was conducted in clean home cages. Briefly, mice were individually placed into a clean cage and left to habituate for 15 min. Then, a social stimulus mouse (5-week-old male C57Bl/6J) was placed into the cage and the experimental mouse was allowed to investigate the stimulus mouse for 10 min. The duration of aspects of social behaviour: head sniffs, anogenital sniffs, and chasing the stimulus mouse were manually scored by an experienced experimenter. Additionally, the duration of grooming was scored.

*Data analysis:* Mouse behavioural data were analysed using GraphPad Prism 8.0 (GraphPad Software, San Diego, USA). Unless described otherwise, data were analysed using one- or two-way ANOVA, with repeated measures were appropriate, followed by LSD *post hoc* tests.

## REFERENCES

1   Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747-753, doi:10.1038/nature08494 (2009).

2   Lee, P. H. *et al.* Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **179**, 1469-1482.e1411, doi:https://doi.org/10.1016/j.cell.2019.11.020 (2019).

3   Sullivan, P. F. *et al.* Psychiatric Genomics: An Update and an Agenda. *The American journal of psychiatry* **175**, 15-27, doi:10.1176/appi.ajp.2017.17030283 (2018).

4  Coleman, J. R. I., Gaspar, H. A., Bryois, J. & Breen, G. The Genetics of the Mood Disorder Spectrum: Genome-wide Association Analyses of More Than 185,000 Cases and 439,000 Controls. *Biological psychiatry*, doi:10.1016/j.biopsych.2019.10.015 (2019).

5  Davies, M. N. *et al.* Generalised Anxiety Disorder--A Twin Study of Genetic Architecture, Genome-Wide Association and Differential Gene Expression. *PloS one* **10**, e0134865, doi:10.1371/journal.pone.0134865 (2015).

6  Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* **50**, 668-681, doi:10.1038/s41588-018-0090-3 (2018).

7  Nagel, M. *et al.* Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nat Genet* **50**, 920-927, doi:10.1038/s41588-018-0151-7 (2018).

8  Fernandez-Castillo, N. *et al.* RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur Neuropsychopharmacol* **30**, 44-55, doi:10.1016/j.euroneuro.2017.11.012 (2020).

9  Bacchelli, E. *et al.* An integrated analysis of rare CNV and exome variation in Autism Spectrum Disorder using the Infinium PsychArray. *Sci Rep* **10**, 3198, doi:10.1038/s41598-020-59922-3 10.1038/s41598-020-59922-3 [pii] (2020).

10 Turner, T. N. *et al.* Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. *Am J Hum Genet* **98**, 58-74, doi:10.1016/j.ajhg.2015.11.023 (2016).

11 Griswold, A. J. *et al.* Targeted massively parallel sequencing of autism spectrum disorder-associated genes in a case control cohort reveals rare loss-of-function risk variants. *Mol Autism* **6**, 43, doi:10.1186/s13229-015-0034-z (2015).

12 Kanduri, C. *et al.* The landscape of copy number variations in Finnish families with autism spectrum disorders. *Autism Res* **9**, 9-16, doi:10.1002/aur.1502 (2016).

13 Zhao, W. W. Intragenic deletion of RBFOX1 associated with neurodevelopmental/neuropsychiatric disorders and possibly other clinical presentations. *Mol Cytogenet* **6**, 26, doi:10.1186/1755-8166-6-26 (2013).

14 Zhang-James, Y. *et al.* An integrated analysis of genes and functional pathways for aggression in human and rodent models. *Molecular Psychiatry* **24**, 1655-1667, doi:10.1038/s41380-018-0068-7 (2019).

15 Richter, J. *et al.* Dynamics of defensive reactivity in patients with panic disorder and agoraphobia: implications for the etiology of panic disorder. *Biological psychiatry* **72**, 512-520, doi:10.1016/j.biopsych.2012.03.035 (2012).

16 Vogt, B. A. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat Rev Neurosci* **6**, 533-544, doi:10.1038/nrn1704 (2005).

17 Laird, A. R. *et al.* Behavioral interpretations of intrinsic connectivity networks. *J Cogn Neurosci* **23**, 4022-4037, doi:10.1162/jocn_a_00077 (2011).

18  Fullana, M. A. *et al.* Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol Psychiatry* **21**, 500-508, doi:10.1038/mp.2015.88 (2016).

19  Yucel, M. *et al.* Anterior cingulate dysfunction: implications for psychiatric disorders? *J Psychiatry Neurosci* **28**, 350-354 (2003).

20  Badre, D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn Sci* **12**, 193-200, doi:10.1016/j.tics.2008.02.004 (2008).

21  Swick, D., Ashley, V. & Turken, A. U. Left inferior frontal gyrus is critical for response inhibition. *BMC Neurosci* **9**, 102, doi:10.1186/1471-2202-9-102 (2008).

22  Perach-Barzilay, N. *et al.* Asymmetry in the dorsolateral prefrontal cortex and aggressive behavior: a continuous theta-burst magnetic stimulation study. *Soc Neurosci* **8**, 178-188, doi:10.1080/17470919.2012.720602 (2013).

23  Alia-Klein, N. *et al.* Trait anger modulates neural activity in the fronto-parietal attention network. *PLoS One* **13**, e0194444, doi:10.1371/journal.pone.0194444 (2018).

24  Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn Sci* **15**, 85-93, doi:10.1016/j.tics.2010.11.004 (2011).

25  Milad, M. R. *et al.* A role for the human dorsal anterior cingulate cortex in fear expression. *Biological psychiatry* **62**, 1191-1194, doi:10.1016/j.biopsych.2007.04.032 (2007).

26  Milad, M. R. & Quirk, G. J. Fear extinction as a model for translational neuroscience: ten years of progress. *Annual review of psychology* **63**, 129-151, doi:10.1146/annurev.psych.121208.131631 (2012).

27  Sehlmeyer, C. *et al.* Human fear conditioning and extinction in neuroimaging: a systematic review. *PloS one* **4**, e5865, doi:10.1371/journal.pone.0005865 (2009).

28  Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. Behavioural phenotyping assays for mouse models of autism. *Nat Rev Neurosci* **11**, 490-502, doi:10.1038/nrn2851 (2010).

29  Jacko, M. *et al.* Rbfox Splicing Factors Promote Neuronal Maturation and Axon Initial Segment Assembly. *Neuron* **97**, 853-868 e856, doi:10.1016/j.neuron.2018.01.020 (2018).

30  Fogel, B. L. *et al.* RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum Mol Genet* **21**, 4171-4186, doi:10.1093/hmg/dds240 (2012).

31  Lee, J. A. *et al.* Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. *Neuron* **89**, 113-128, doi:10.1016/j.neuron.2015.11.025 (2016).

32  Li, Y. I., Sanchez-Pulido, L., Haerty, W. & Ponting, C. P. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res* **25**, 1-13, doi:10.1101/gr.181990.114 (2015).

33  Wamsley, B. *et al.* Rbfox1 Mediates Cell-type-Specific Splicing in Cortical Interneurons. *Neuron* **100**, 846-859 e847, doi:10.1016/j.neuron.2018.09.026 (2018).

34  Vuong, C. K. *et al.* Rbfox1 Regulates Synaptic Transmission through the Inhibitory Neuron-Specific vSNARE Vamp1. *Neuron* **98**, 127-141 e127, doi:10.1016/j.neuron.2018.03.008 (2018).

35  Gehman, L. T. *et al.* The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat Genet* **43**, 706-711, doi:10.1038/ng.841 (2011).

36 Ferguson, B. R. & Gao, W. J. PV Interneurons: Critical Regulators of E/I Balance for Prefrontal Cortex-Dependent Behavior and Psychiatric Disorders. *Front Neural Circuits* **12**, 37, doi:10.3389/fncir.2018.00037 (2018).

37 Tomassoni-Ardori, F. *et al.* Rbfox1 up-regulation impairs BDNF-dependent hippocampal LTP by dysregulating TrkB isoform expression levels. *Elife* **8**, doi:10.7554/eLife.49673 (2019).

38 Pezawas, L. *et al.* 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nat Neurosci* **8**, 828-834, doi:10.1038/nn1463 (2005).

39 Schweiger, J. I. *et al.* Effects of BDNF Val(66)Met genotype and schizophrenia familial risk on a neural functional network for cognitive control in humans. *Neuropsychopharmacology* **44**, 590-597, doi:10.1038/s41386-018-0248-9 (2019).

40 Tost, H. *et al.* Effects of the BDNF Val66Met polymorphism on white matter microstructure in healthy adults. *Neuropsychopharmacology* **38**, 525-532, doi:10.1038/npp.2012.214 (2013).

41 Meyer-Lindenberg, A. *et al.* Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat Neurosci* **8**, 594-596, doi:10.1038/nn1438 (2005).

42 Braun, U. *et al.* Dynamic brain network reconfiguration as a potential schizophrenia genetic risk mechanism modulated by NMDA receptor function. *Proc Natl Acad Sci U S A* **113**, 12568-12573, doi:10.1073/pnas.1608819113 (2016).

43 Schneider, M. *et al.* Altered DLPFC-Hippocampus Connectivity During Working Memory: Independent Replication and Disorder Specificity of a Putative Genetic Risk Phenotype for Schizophrenia. *Schizophr Bull* **43**, 1114-1122, doi:10.1093/schbul/sbx001 (2017).

44 Cao, H. *et al.* Altered Functional Subnetwork During Emotional Face Processing: A Potential Intermediate Phenotype for Schizophrenia. *JAMA Psychiatry* **73**, 598-605, doi:10.1001/jamapsychiatry.2016.0161 (2016).

45 Grimm, O. *et al.* Striatal response to reward anticipation: evidence for a systems-level intermediate phenotype for schizophrenia. *JAMA Psychiatry* **71**, 531-539, doi:10.1001/jamapsychiatry.2014.9 (2014).

46 Harneit, A. *et al.* MAOA-VNTR genotype affects structural and functional connectivity in distributed brain networks. *Hum Brain Mapp* **40**, 5202-5212, doi:10.1002/hbm.24766 (2019).

47 Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F. & Weinberger, D. R. The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage* **17**, 317-323, doi:10.1006/nimg.2002.1179 (2002).

48 Meyer-Lindenberg, A. *et al.* Neural mechanisms of genetic risk for impulsivity and violence in humans. *Proc Natl Acad Sci U S A* **103**, 6269-6274, doi:10.1073/pnas.0511311103 (2006).

49 Tzourio-Mazoyer, N. *et al.* Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273-289, doi:10.1006/nimg.2001.0978 (2002).

50 Bush, G., Luu, P. & Posner, M. I. Cognitive and emotional influences in anterior cingulate cortex. *Trends in cognitive sciences* **4**, 215-222 (2000).

51 Kircher, T. *et al.* Effect of cognitive-behavioral therapy on neural correlates of fear conditioning in panic disorder. *Biological psychiatry* **73**, 93-101, doi:10.1016/j.biopsych.2012.07.026 (2013).

52 Lueken, U. *et al.* Neural substrates of treatment response to cognitive-behavioral therapy in panic disorder with agoraphobia. *The American journal of psychiatry* **170**, 1345-1355, doi:10.1176/appi.ajp.2013.12111484 (2013).

53 Yang, Y. *et al.* Effect of CBT on Biased Semantic Network in Panic Disorder: A Multicenter fMRI Study Using Semantic Priming. *The American journal of psychiatry* **177**, 254-264, doi:10.1176/appi.ajp.2019.19020202 (2020).

54 Horner, A. E. *et al.* The touchscreen operant platform for testing learning and memory in rats and mice. *Nat Protoc* **8**, 1961-1984, doi:10.1038/nprot.2013.122 (2013).

# SUPPLEMENTARY MATERIAL

1. **SUPPLEMENTARY FIGURES**

2. **SUPPLEMENTARY TABLES**

# 1. SUPPLEMENTARY FIGURES

**A** CNVs in patients with other psychiatric disorders

ADHD   BIP   OCD   TS   CD

**B** CNVs in controls
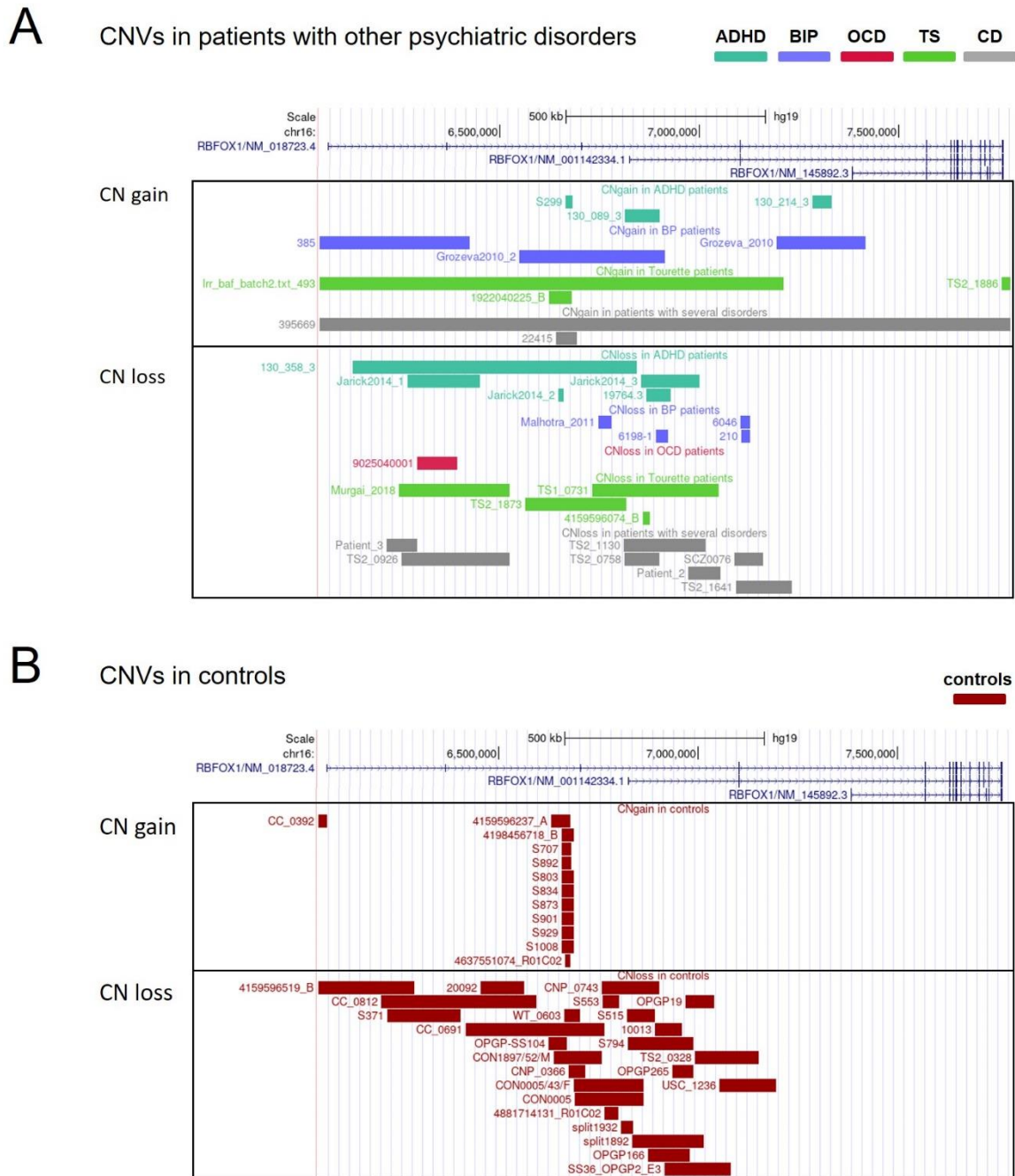
controls

**Supplementary Figure 1**. Copy number variants in *RBFOX1* identified in A) patients with other psychiatric disorders, B) controls described in the 18 papers used for the burden test analysis.

**A** Peak regions enriched in gains and losses
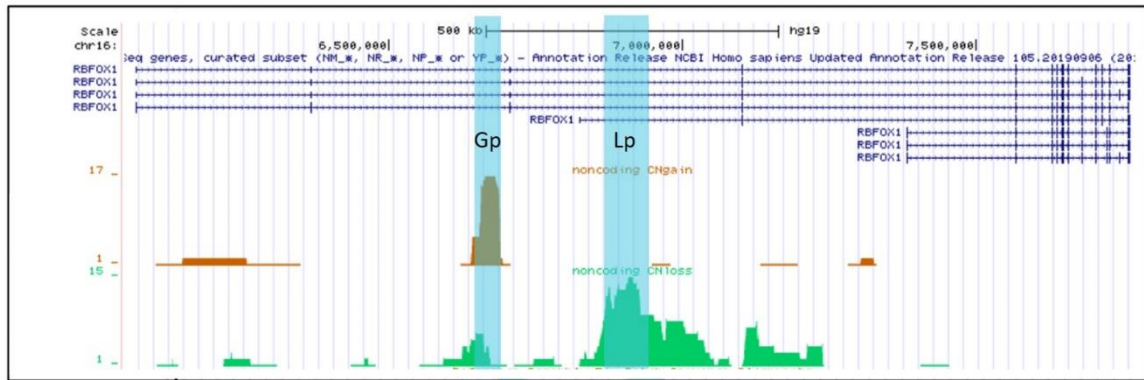
**B** Gain peak
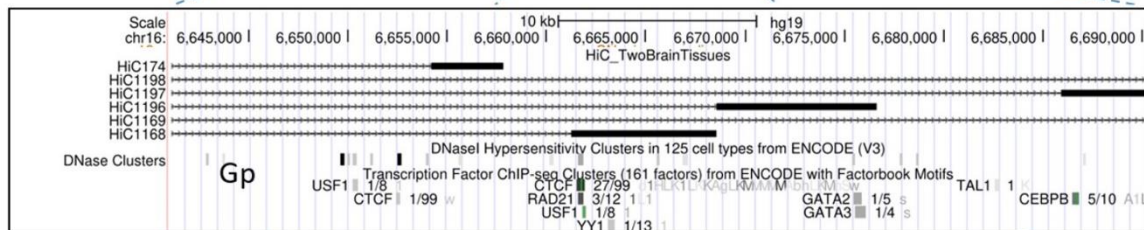
**C** Loss peak
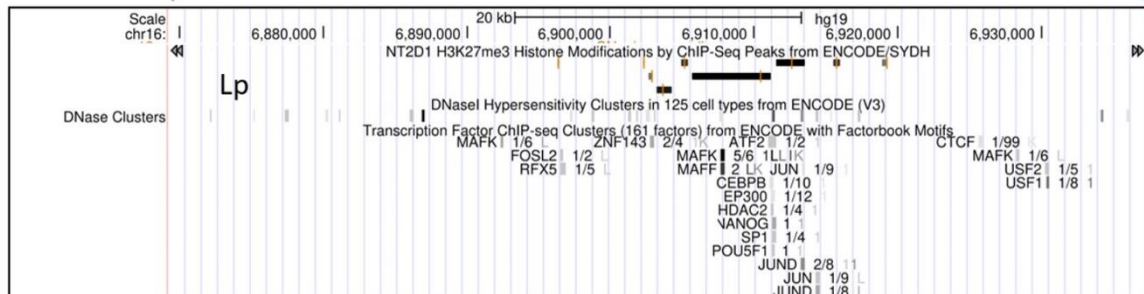
**Supplementary Figure 2**. Distribution of non-coding copy number variants (CNVs) in *RBFOX1* in patients with psychiatric conditions and overlap with transcriptional regulatory elements. a) Distribution of non-coding CNVs in *RBFOX1* in patients. Peak regions enriched in CNVs are highlighted in blue. Gp, gain peak, corresponds to the region where more CN gains are concentrated; Lp, loss peak, corresponds to the region where more CN losses are concentrated. b) Distribution of ChIP-seq signatures from ENCODE present in the Lp region. c) Distribution of ChIP-seq signatures from ENCODE present in the Gp region.
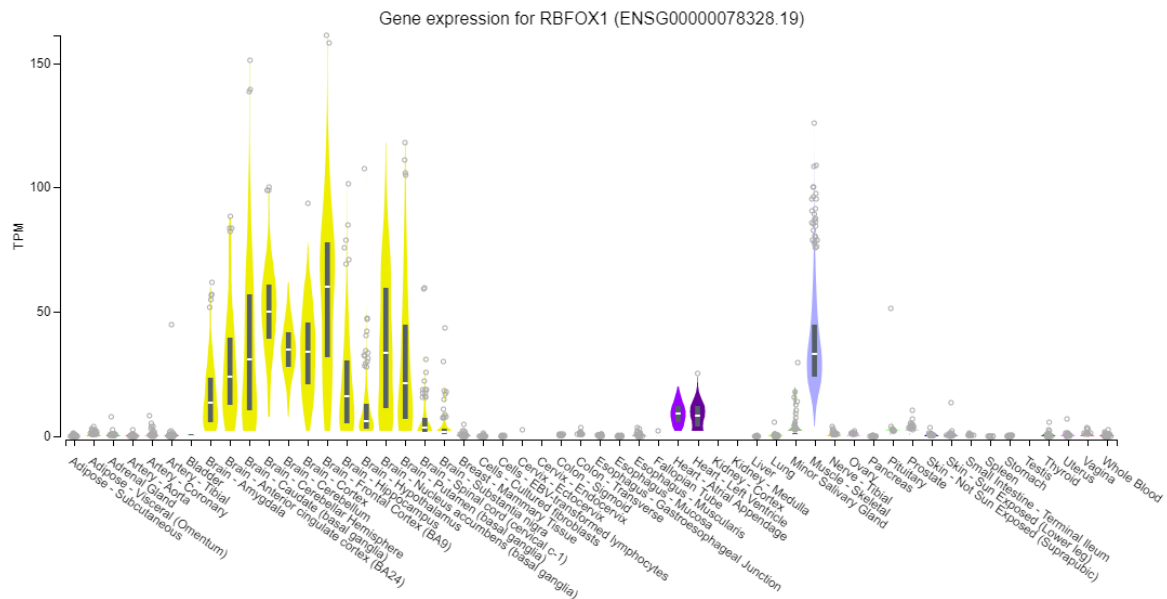
**Supplementary Figure 3**. *RBFOX1* expression in different tissues from human samples (GTEx database).
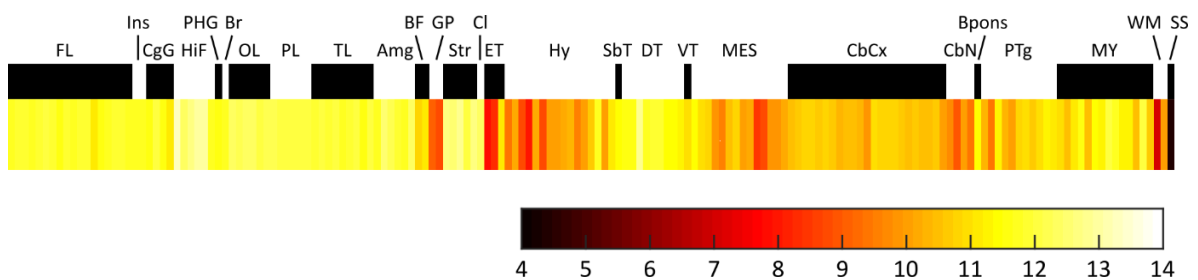


**Supplementary Figure 4**. *RBFOX1* gene expression in different brain regions of the human brain (FL, frontal lobe; Ins, Insula; CgG, Cingulate Gyrus; HiF, Hippocampal Formation; PHG, parahippocampal gyrus; Br, piriform cortex; OL, occipital lobe; PL, parietal lobe; TL, temporal lobe; Amg, amygdala; BF, basal forebrain; GP, globus pallidus; Str, striatum; Cl, claustrum; ET, epithalamus; Hy, hypothalamus; SbT, subthalamus; DT, dorsal thalamus; VT, ventral thalamus; MES, mesencephalon; CbCx, cerebellar cortex; CbN, cerebellar nuclei; Bpons, basal part of pons; PTg, pontine tegmentum; MY, myencephalon; WM, white matter; SS, sulci & spaces). Data was obtained from the Allen Human Brain Atlas (http://human.brain-map.org/) (Hawrylycz et al. 2012) and depicted as the average from two different probes.

**RBFOX1 Expression Changes**
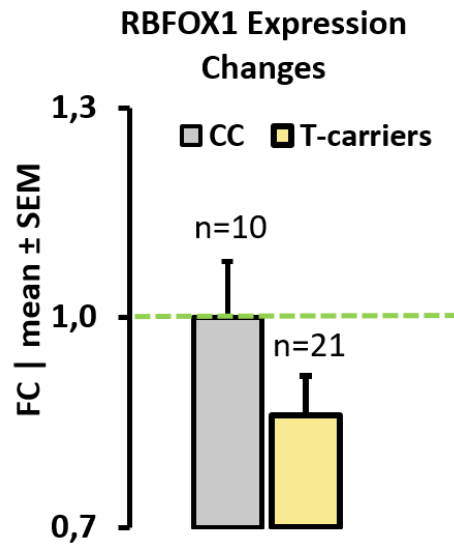
**Supplementary Figure 5**. Relative RBFOX1 protein expression in human post-mortem frontal cortex in *rs6500744* CC genotype vs T-allele carriers.



a) marble burying test

b) pairwise discrimination task
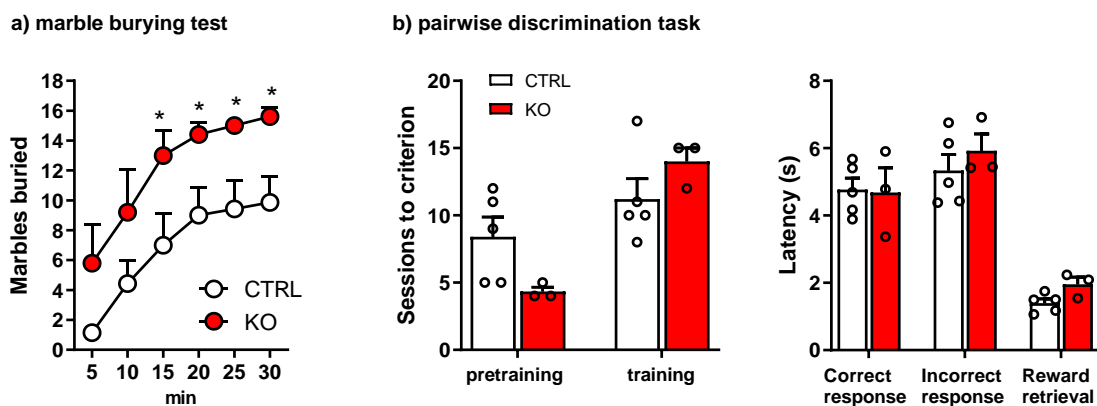
**Supplementary Figure 6**. Behavioural effects of neuron-specific deletion of *Rbfox1* in mice. **A,** *Rbfox1*-KO buried more marbles than CTRL in the marble-burying test but this appeared to be due to excessive digging and displacement of bedding rather than anxiety (Supplementary Video 1). **B,** there were no differences between CTRL and KO in the acquisition of the touchscreen pairwise discrimination task (as measured by the number of sessions completed before reaching the criterion); additionally, similar response latencies to stimuli and reward retrieval suggest that CTRL and KO did not differ in motivation to perform the task.

## 2. SUPPLEMANTARY TABLES (

**Supplementary Table 1.** Description of samples used for analyses of common genetic variants (summary statistics from GWAS meta-analysis)

| GWAS | Abbreviation | Participants | Reference |
|---|---|---|---|
| Attention-Deficit Hyperactivity Disorder | ADHD | 19,099 Ca + 34,194 Co | Demontis et al., 2019 |
| Aggression | AGG | 87,485 individuals | Ip et al., 2019 (Biorxvs) |
| Autism Spectrum Disorder | ASD | 18,382 Ca + 27,969 Co | Grove et al., 2019 |
| Anorexia | ANO | 16,992 Ca + 55,525 Co | Watson et al., 2019 |
| Anxiety | ANX | 12,655 Ca + 19,255 Co | Meier et al., 2019 |
| Bipolar Disorder | BIP | 20,352 Ca + 31,358 Co | Stahl et al., 2019 |
| Major Depressive Disorder | MDD | 135,458 Ca + 344,901 Co | Wray et al., 2018 |
| Obsessive-Compulsive Disorder | OCD | 1,773 Ca + 6,122 Co + 915 trios | Arnold et al., 2018 |
| Risk tolerance behavior | RT | 975,353 individuals | Linner et al., 2019 |
| Schizophrenia | SCZ | 67,280 Ca + 86,912 Co | Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2020 |
| Tourette's Syndrome | TS | 4,819 Ca + 9,488 Co | Yu et al., 2019 |
| Cross-Disorder meta-analysis | CD-MA | 232,964 Ca + 494,162 Co | Lee et al. 2019 |

Ca = Cases; Co = controls. Summary statistics from these GWAS meta-analyses were obtained from the Psychiatric Genomics Consortium (PGC; https://www.med.unc.edu/pgc/download-results/), Integrative Psychiatric Research Consortium (iPSYCH; https://ipsych.dk/en/research/downloads/), UK Biobank (https://www.ukbiobank.ac.uk/) and 23andMe (https://research.23andme.com/) or authors' request.

**Supplementary Table 2**. Common genetic risk variants (SNPs) in RBFOX1 gene showing suggestive associations with psychiatric conditions

| Disorder | SNP | P |
|---|---|---|
| SCZ | rs8056990 | 6.50E-06 |
| SCZ | rs1906061 | 6.80E-06 |
| SCZ | rs12448737m | 7.39E-06 |
| SCZ | rs8050094 | 8.59E-06 |
| SCZ | rs8049954 | 8.64E-06 |
| SCZ | rs6500742 | 8.71E-06 |
| RT | rs4479249 | **1.77E-08** |
| RT | rs6500948 | **2.13E-08** |
| RT | rs4350587 | **2.96E-08** |
| RT | rs4786996 | **4.75E-08** |
| RT | rs8046401 | 5.36E-08 |
| RT | rs9940929 | 6.18E-08 |
| RT | rs4238877 | 1.09E-07 |
| RT | rs4337311 | 1.15E-07 |
| RT | rs4290489 | 1.35E-07 |
| RT | rs2178721 | 1.54E-07 |
| RT | rs7200150 | 1.84E-07 |
| RT | rs4787000 | 1.85E-07 |
| RT | rs12925090 | 2.34E-07 |
| RT | rs12925403 | 2.53E-07 |
| RT | rs4494563 | 3.31E-07 |
| RT | rs6500947 | 7.23E-07 |
| RT | rs2191131 | 1.14E-06 |
| RT | rs2191130 | 1.15E-06 |
| RT | rs7204945 | 1.64E-06 |
| RT | rs7202295 | 1.73E-06 |
| RT | rs6500946 | 1.96E-06 |
| RT | rs4511555 | 2.25E-06 |
| RT | rs4473206 | 2.30E-06 |
| RT | rs5009028 | 2.30E-06 |
| RT | rs4786997 | 2.32E-06 |
| RT | rs34857835 | 2.35E-06 |
| RT | rs35378747 | 2.40E-06 |
| RT | rs59271741 | 2.49E-06 |
| RT | rs4786998 | 2.51E-06 |
| RT | rs4411517 | 2.67E-06 |
| RT | rs4497710 | 2.78E-06 |
| RT | rs12445831 | 2.87E-06 |
| RT | rs4360957 | 2.93E-06 |
| RT | rs4381618 | 2.99E-06 |
| RT | rs12447663 | 3.05E-06 |
| RT | rs11862622 | 3.13E-06 |
| RT | rs12445208 | 3.19E-06 |
| RT | rs11866790 | 3.20E-06 |
| RT | rs11862619 | 3.22E-06 |
| RT | rs17562208 | 3.29E-06 |
| RT | rs11861673 | 3.32E-06 |
| RT | rs5009029 | 3.43E-06 |
| RT | rs11861497 | 3.49E-06 |
| RT | rs4787050 | 3.51E-06 |

| | | |
|---|---|---|
| **RT** | rs17143433 | 3.56E-06 |
| **RT** | rs4787048 | 3.90E-06 |
| **RT** | rs3785232 | 3.93E-06 |
| **RT** | rs7189741 | 4.17E-06 |
| **RT** | rs7185128 | 4.19E-06 |
| **RT** | rs4786995 | 4.20E-06 |
| **RT** | rs11645478 | 5.11E-06 |
| **RT** | rs12924980 | 5.23E-06 |
| **RT** | rs11077169 | 6.18E-06 |
| **RT** | rs4787008 | 6.70E-06 |
| **RT** | rs1935397 | 7.75E-06 |
| **RT** | rs2534760 | 8.49E-06 |
| **CD-MA** | rs7193263 | **1.37E-12** |
| **CD-MA** | rs8063603 | **1.88E-12** |
| **CD-MA** | rs10852676 | **1.02E-11** |
| **CD-MA** | rs6500768 | **2.94E-11** |
| **CD-MA** | rs716508 | **4.42E-11** |
| **CD-MA** | rs7188257 | **4.96E-11** |
| **CD-MA** | rs61547418 | **9.83E-11** |
| **CD-MA** | rs6500770 | **1.12E-10** |
| **CD-MA** | rs4786850 | **1.15E-10** |
| **CD-MA** | rs1861188 | **1.31E-10** |
| **CD-MA** | rs113726301 | **1.33E-10** |
| **CD-MA** | rs12448420 | **1.50E-10** |
| **CD-MA** | rs72778205 | **1.50E-10** |
| **CD-MA** | rs57035629 | **1.81E-10** |
| **CD-MA** | rs111429920 | **2.10E-10** |
| **CD-MA** | rs113039090 | **2.21E-10** |
| **CD-MA** | rs60856919 | **2.64E-10** |
| **CD-MA** | rs7187481 | **2.87E-10** |
| **CD-MA** | rs111841592 | **3.06E-10** |
| **CD-MA** | rs188088642 | **3.15E-10** |
| **CD-MA** | rs7188476 | **3.46E-10** |
| **CD-MA** | rs11077022 | **3.61E-10** |
| **CD-MA** | rs60148766 | **4.99E-10** |
| **CD-MA** | rs10852675 | **5.26E-10** |
| **CD-MA** | rs8050261 | **9.77E-10** |
| **CD-MA** | rs8051084 | **1.03E-09** |
| **CD-MA** | rs8050918 | **1.08E-09** |
| **CD-MA** | rs10852673 | **1.11E-09** |
| **CD-MA** | rs12448139 | **3.26E-09** |
| **CD-MA** | rs17139688 | **3.58E-09** |
| **CD-MA** | rs10500337 | **3.91E-09** |
| **CD-MA** | rs12927291 | **4.18E-09** |
| **CD-MA** | rs12928387 | **4.87E-09** |
| **CD-MA** | rs56354361 | **5.13E-09** |
| **CD-MA** | rs8054572 | **5.26E-09** |
| **CD-MA** | rs10500338 | **6.64E-09** |
| **CD-MA** | rs6500765 | **9.96E-09** |
| **CD-MA** | rs1420025 | **1.14E-08** |
| **CD-MA** | rs6500769 | **1.44E-08** |
| **CD-MA** | rs8049123 | **1.88E-08** |
| **CD-MA** | rs11077204 | **4.59E-08** |
| **CD-MA** | rs3785236 | **4.63E-08** |

| | | |
|---|---|---|
| **CD-MA** | rs13338644 | 9.04E-08 |
| **CD-MA** | rs7202627 | 9.37E-08 |
| **CD-MA** | rs7199065 | 9.79E-08 |
| **CD-MA** | rs3785234 | 1.05E-07 |
| **CD-MA** | rs11077024 | 1.10E-07 |
| **CD-MA** | rs17221054 | 1.12E-07 |
| **CD-MA** | rs11077206 | 1.17E-07 |
| **CD-MA** | rs7198928 | 1.49E-07 |
| **CD-MA** | rs35851985 | 1.66E-07 |
| **CD-MA** | rs57256407 | 1.67E-07 |
| **CD-MA** | rs11077203 | 1.80E-07 |
| **CD-MA** | rs9925434 | 2.08E-07 |
| **CD-MA** | rs58330268 | 2.15E-07 |
| **CD-MA** | rs4786091 | 2.23E-07 |
| **CD-MA** | rs7186816 | 2.26E-07 |
| **CD-MA** | rs11640652 | 2.27E-07 |
| **CD-MA** | rs7204695 | 2.54E-07 |
| **CD-MA** | rs1978314 | 2.65E-07 |
| **CD-MA** | rs4274443 | 2.85E-07 |
| **CD-MA** | rs3785235 | 2.89E-07 |
| **CD-MA** | rs57621245 | 2.92E-07 |
| **CD-MA** | rs1362318 | 2.93E-07 |
| **CD-MA** | rs17221152 | 3.02E-07 |
| **CD-MA** | rs1420036 | 3.05E-07 |
| **CD-MA** | rs7192025 | 3.20E-07 |
| **CD-MA** | rs17139767 | 3.22E-07 |
| **CD-MA** | rs933479 | 3.22E-07 |
| **CD-MA** | rs55997507 | 3.25E-07 |
| **CD-MA** | rs1978316 | 3.31E-07 |
| **CD-MA** | rs7202054 | 3.36E-07 |
| **CD-MA** | rs4255786 | 3.53E-07 |
| **CD-MA** | rs4255787 | 3.53E-07 |
| **CD-MA** | rs55680138 | 3.59E-07 |
| **CD-MA** | rs67861918 | 3.65E-07 |
| **CD-MA** | rs7191889 | 3.65E-07 |
| **CD-MA** | rs17220445 | 3.70E-07 |
| **CD-MA** | rs7191200 | 3.78E-07 |
| **CD-MA** | rs17220529 | 3.93E-07 |
| **CD-MA** | rs1344470 | 4.18E-07 |
| **CD-MA** | rs4786848 | 4.23E-07 |
| **CD-MA** | rs113752785 | 4.32E-07 |
| **CD-MA** | rs11863506 | 4.33E-07 |
| **CD-MA** | rs2880916 | 4.39E-07 |
| **CD-MA** | rs66495252 | 4.44E-07 |
| **CD-MA** | rs1344471 | 4.46E-07 |
| **CD-MA** | rs7190951 | 4.48E-07 |
| **CD-MA** | rs1344472 | 4.50E-07 |
| **CD-MA** | rs67162703 | 4.66E-07 |
| **CD-MA** | rs17220612 | 4.78E-07 |
| **CD-MA** | rs73521157 | 4.84E-07 |
| **CD-MA** | rs72776497 | 4.86E-07 |
| **CD-MA** | rs11867043 | 5.02E-07 |
| **CD-MA** | rs716509 | 5.03E-07 |
| **CD-MA** | rs1420037 | 5.13E-07 |

| | | |
|---|---|---|
| **CD-MA** | rs73521156 | 5.24E-07 |
| **CD-MA** | rs67023896 | 5.36E-07 |
| **CD-MA** | rs4786092 | 5.42E-07 |
| **CD-MA** | rs111420470 | 5.54E-07 |
| **CD-MA** | rs3785238 | 5.66E-07 |
| **CD-MA** | rs17819872 | 5.76E-07 |
| **CD-MA** | rs13336122 | 5.89E-07 |
| **CD-MA** | rs62016049 | 5.95E-07 |
| **CD-MA** | rs17219976 | 6.50E-07 |
| **CD-MA** | rs4786841 | 6.64E-07 |
| **CD-MA** | rs17819914 | 6.72E-07 |
| **CD-MA** | rs66589142 | 6.93E-07 |
| **CD-MA** | rs11077016 | 6.95E-07 |
| **CD-MA** | rs1978317 | 7.02E-07 |
| **CD-MA** | rs67962445 | 7.03E-07 |
| **CD-MA** | rs7196688 | 7.15E-07 |
| **CD-MA** | rs68116725 | 7.19E-07 |
| **CD-MA** | rs73521144 | 7.90E-07 |
| **CD-MA** | rs9935068 | 8.35E-07 |
| **CD-MA** | rs56193200 | 8.35E-07 |
| **CD-MA** | rs1344469 | 8.55E-07 |
| **CD-MA** | rs72776482 | 8.59E-07 |
| **CD-MA** | rs73521143 | 8.86E-07 |
| **CD-MA** | rs67239940 | 8.97E-07 |
| **CD-MA** | rs72776495 | 8.98E-07 |
| **CD-MA** | rs1547539 | 9.01E-07 |
| **CD-MA** | rs62016050 | 9.48E-07 |
| **CD-MA** | rs17819962 | 9.53E-07 |
| **CD-MA** | rs72776485 | 9.61E-07 |
| **CD-MA** | rs12446403 | 1.09E-06 |
| **CD-MA** | rs17219189 | 1.21E-06 |
| **CD-MA** | rs67577697 | 1.26E-06 |
| **CD-MA** | rs67347747 | 1.32E-06 |
| **CD-MA** | rs12446113 | 1.39E-06 |
| **CD-MA** | rs9929993 | 1.39E-06 |
| **CD-MA** | rs1978315 | 1.48E-06 |
| **CD-MA** | rs72776412 | 1.58E-06 |
| **CD-MA** | rs67517937 | 1.84E-06 |
| **CD-MA** | rs67896120 | 1.85E-06 |
| **CD-MA** | rs7187868 | 1.88E-06 |
| **CD-MA** | rs2534753 | 1.98E-06 |
| **CD-MA** | rs10500339 | 2.00E-06 |
| **CD-MA** | rs72776415 | 2.01E-06 |
| **CD-MA** | rs61563561 | 2.07E-06 |
| **CD-MA** | rs77752707 | 2.36E-06 |
| **CD-MA** | rs11077202 | 2.36E-06 |
| **CD-MA** | rs67607023 | 2.64E-06 |
| **CD-MA** | rs12444690 | 2.85E-06 |
| **CD-MA** | rs2160166 | 3.46E-06 |
| **CD-MA** | rs4787049 | 3.61E-06 |
| **CD-MA** | rs56308757 | 3.73E-06 |
| **CD-MA** | rs9935836 | 4.35E-06 |
| **CD-MA** | rs55793113 | 6.11E-06 |
| **CD-MA** | rs4616299 | 6.13E-06 |

| CD-MA | rs11640969 | 6.42E-06 |
| CD-MA | rs78150755 | 6.55E-06 |
| CD-MA | rs77630439 | 6.55E-06 |
| CD-MA | rs11640647 | 6.73E-06 |
| CD-MA | rs1640880 | 8.10E-06 |
| MDD | rs7198928 | **4.45E-11** |
| MDD | rs11077206 | **4.45E-11** |
| MDD | rs3785234 | **4.45E-11** |
| MDD | rs9925434 | **4.45E-11** |
| MDD | rs3785235 | **6.44E-11** |
| MDD | rs11077203 | **7.74E-11** |
| MDD | rs35851985 | **9.29E-11** |
| MDD | rs3785238 | **1.22E-10** |
| MDD | rs7192025 | **1.34E-10** |
| MDD | rs55997507 | **2.74E-10** |
| MDD | rs7193263 | **4.33E-10** |
| MDD | rs8063603 | **6.04E-10** |
| MDD | rs11640652 | **9.37E-10** |
| MDD | rs7202054 | **9.37E-10** |
| MDD | rs7191889 | **1.11E-09** |
| MDD | rs7191200 | **1.32E-09** |
| MDD | rs7190951 | **1.32E-09** |
| MDD | rs9929993 | **1.35E-09** |
| MDD | rs7196688 | **2.61E-09** |
| MDD | rs3785232 | **2.61E-09** |
| MDD | rs2191130 | **3.06E-09** |
| MDD | rs2191131 | **3.06E-09** |
| MDD | rs10852687 | **4.15E-09** |
| MDD | rs7204945 | **4.20E-09** |
| MDD | rs3785236 | **8.38E-09** |
| MDD | rs4787050 | **9.17E-09** |
| MDD | rs11077204 | **9.67E-09** |
| MDD | rs7184911 | **1.16E-08** |
| MDD | rs11646221 | **2.19E-08** |
| MDD | rs10852673 | **2.90E-08** |
| MDD | rs7195278 | **3.01E-08** |
| MDD | rs4787048 | **3.58E-08** |
| MDD | rs1861188 | **3.58E-08** |
| MDD | rs10852676 | **3.58E-08** |
| MDD | rs716508 | **3.58E-08** |
| MDD | rs11640647 | **4.12E-08** |
| MDD | rs11640969 | **4.12E-08** |
| MDD | rs6500768 | **4.81E-08** |
| MDD | rs9935836 | 5.12E-08 |
| MDD | rs7188257 | 5.17E-08 |
| MDD | rs716983 | 5.62E-08 |
| MDD | rs4787049 | 5.79E-08 |
| MDD | rs12444690 | 6.55E-08 |
| MDD | rs11077022 | 7.46E-08 |
| MDD | rs7186834 | 1.04E-07 |
| MDD | rs8051084 | 1.05E-07 |
| MDD | rs8050918 | 1.21E-07 |
| MDD | rs3785237 | 1.34E-07 |
| MDD | rs8050261 | 1.39E-07 |

| | | |
|---|---|---|
| **MDD** | rs8054572 | 1.44E-07 |
| **MDD** | rs12927291 | 1.65E-07 |
| **MDD** | rs6500769 | 1.71E-07 |
| **MDD** | rs2191129 | 2.12E-07 |
| **MDD** | rs12928387 | 2.16E-07 |
| **MDD** | rs10500338 | 2.16E-07 |
| **MDD** | rs12448139 | 2.16E-07 |
| **MDD** | rs13338644 | 2.33E-07 |
| **MDD** | rs1547539 | 2.54E-07 |
| **MDD** | rs7187481 | 2.75E-07 |
| **MDD** | rs11077207 | 2.87E-07 |
| **MDD** | rs10852675 | 3.08E-07 |
| **MDD** | rs4786850 | 3.14E-07 |
| **MDD** | rs10500337 | 3.21E-07 |
| **MDD** | rs10492762 | 3.33E-07 |
| **MDD** | rs58775805 | 3.57E-07 |
| **MDD** | rs17139688 | 3.66E-07 |
| **MDD** | rs11641267 | 3.87E-07 |
| **MDD** | rs3785229 | 3.87E-07 |
| **MDD** | rs6500770 | 4.11E-07 |
| **MDD** | rs56354361 | 4.17E-07 |
| **MDD** | rs11640510 | 4.49E-07 |
| **MDD** | rs8049123 | 4.74E-07 |
| **MDD** | rs17685565 | 5.20E-07 |
| **MDD** | rs34518736 | 5.20E-07 |
| **MDD** | rs6500765 | 5.40E-07 |
| **MDD** | rs12923556 | 6.03E-07 |
| **MDD** | rs7188476 | 6.12E-07 |
| **MDD** | rs12923795 | 9.33E-07 |
| **MDD** | rs1420025 | 1.16E-06 |
| **MDD** | rs1857952 | 1.22E-06 |
| **MDD** | rs61547418 | 1.43E-06 |
| **MDD** | rs113726301 | 1.61E-06 |
| **MDD** | rs111429920 | 2.29E-06 |
| **MDD** | rs111841592 | 2.58E-06 |
| **MDD** | rs188088642 | 2.58E-06 |
| **MDD** | rs113039090 | 2.89E-06 |
| **MDD** | rs4616299 | 3.18E-06 |
| **MDD** | rs72778205 | 4.56E-06 |
| **MDD** | rs10492829 | 4.64E-06 |
| **MDD** | rs78728828 | 4.64E-06 |
| **MDD** | rs72762903 | 5.10E-06 |
| **MDD** | rs60148766 | 5.11E-06 |
| **MDD** | rs4786872 | 5.57E-06 |
| **MDD** | rs60856919 | 5.72E-06 |
| **MDD** | rs17140337 | 6.41E-06 |
| **MDD** | rs57035629 | 8.42E-06 |

SNP, single nucleotide polymorphism; CD-MA, cross-disorder meta-analysis; MDD, major depressive disorder; SCZ, schizophrenia; RT, risk tolerance behavior. In bold, SNPs reaching genome-wide significance (p<5.0E-08).

**Supplementary table 3.** Enrichment of *RBFOX1* target genes among the significant genes associated for each psychiatric condition

|  | ADHD | AGG | ANO | ANX | ASD | BIP | MDD | OCD | RT | TS | SCZ | CD-MA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Number of associated genes in the gene-based analysis\*** | 20 | 3 | 41 | 2 | 13 | 1 | 262 | 1 | 279 | 2 | 438 | 266 |
| **Number of *RBFOX1* target genes among the gene-based associated genes** | 4 | 1 | 8 | 1 | 1 | 0 | 42 | 0 | 46 | 1 | 60 | 42 |
| **p-value of the hypergeometric test** | 0.140 | 0.287 | 0.069 | 0.219 | 0.327 | N/A | **0.016** | N/A | **0.010** | 0.219 | **0.042** | **0.019** |

\* Number of genes associated in the GWAS of each psychiatric condition overcoming gene-wide significance (Bonferroni correction; p = 20000/0.05 = 2.5E-06). ADHD, attention-deficit and hyperactivity disorder; AGG, childhood aggression; ANO, anorexia nervosa; ANX, anxiety; ASB, antisocial behaviour; ASD, autism spectrum disorder; BP, bipolar disorder; CD-MA, cross-disorder meta-analysis; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; SCZ, schizophrenia; RT, risk tolerance; TS, Tourette syndrome. Total number of genes in the genome: 20000 genes; 2499 *RBFOX1* target genes (Lee et al. 2016).

**Supplementary Table 4.** CNVs found in RBFOX1 in patients with psychiatric conditions

| Mutation, change | Inher. | Individual | Sex | Coordinates (hg19 build) | Size (bp) | Exonic/Intronic/UTR | Exons and introns affected | Disorder / comorbidity | Family | Study (PMID) |
|---|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | P | 130_358_3 | N/A | chr16:6132598-6843831 | 711,234 | UTR, exonic, intronic | introns 1-3, exons 2-3 (NM_018723) / UTR, intron 1 (NM_01142334) | ADHD | father likely meets criteria for adulthood ADHD (ASRS) | Elia et al., 2010 (19546859) |
| Copy number loss | N/A | S182 | N/A | chr16:6269342-6451865 | 182,523 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | ADHD | N/A | Jarick et al., 2014 (23164820) |
| Copy number loss | N/A | S180 | N/A | chr16:6648152-6660935 | 12783 | intronic | intron 2 (NM_018723) | ADHD | N/A | Jarick et al., 2014 (23164820) |
| Copy number loss | N/A | S169 | N/A | chr16:6855937-7001231 | 145294 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | ADHD | N/A | Jarick et al., 2014 (23164820) |
| Copy number loss | N/A | 19764.3 | F | chr16:6868721-6928434 | 59,714 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | ADHD | N/A | Lionel et al.. 2011 (21832240) |
| Copy number loss | N/A | 44307 | M | chr16:6637427-6661127 | 23,701 | intronic | intron 2 (NM_018723) | ASD | N/A | Prasad et al., 2012 (23275889) |
| Copy number loss | N/A | 68257 | M | chr16:6870651-6926527 | 55,877 | intronic | intron 1 (NM_001142334) / intron 3 (NM_018723) | ASD | N/A | Prasad et al., 2012 (23275889) |
| Copy number loss | N/A | 107204L | M | chr16:7213159-7241549 | 28,391 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Prasad et al., 2012 (23275889) |
| Copy number loss | M | AB74 | N/A | chr16:6666748-6700500 | 33,752 | intronic | intron 2 (NM_018723) | ASD | N/A | Bacchelli et al., 2020 (32081867) |
| Copy number loss | M | AB86 | N/A | chr16:6816914-6929536 | 112,622 | UTR, exonic, intronic | UTR exon 1, intron 1 (NM_001142334) / intron 3 (NM_018723) | ASD | N/A | Bacchelli et al., 2020 (32081867) |
| Copy number loss | P | 6264_3 | M | chr16:6626430-6669592 | 43,163 | intronic | intron 2 (NM_018723) | ASD | Simplex | Pinto et al., 2010 (20531469) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | Ml | 6317_5 | M | chr16:6864528-6927242 | 62,715 | intronic | intron 1 (NM_001142334) / intron 3 (NM_018723) | ASD | Simplex | Pinto et al., 2010 (20531469) |
| Copy number loss | M | 5244_3 | M | chr16:6870628-6929560 | 58,933 | exonic, intronic | intron 1 (NM_001142334) / intron 3 (NM_018723) | ASD | Simplex | Pinto et al., 2010 (20531469) |
| Copy number loss | M | 14146_2440 | F | chr16:7040830-7126732 | 85,903 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, exons 1-2, UTR (NM_001142334) | ASD | Simplex | Pinto et al., 2010 (20531469) |
| Copy number loss | P | N/A | M | chr16:6296188-6309514 | 13,326 | intronic | intron 1 (NM_018723) | ASD, developmental hemiparesis | Simplex, father diagnosed with anxiety | Davis et al., 2012 (22678932) |
| Copy number loss | de novo | AU077504 | F | chr16:6052835-6260815 | 207,982 | exonic, intronic | UTR, exon 1 , intron 1 (NM_018723) | ASD, epilepsy, mental retardation | Simplex | Sebat et al., 2007 (17363630); Martin et al., 2007 (17503474) |
| Copy number loss | P | 7636 | M | chr16:6926261-7102997 | 176,736 | UTR, exonic, intronic | exon 4, introns 3-4 (NM_018723) / introns 3-4, UTR, exons 1-2 (NM_001142334) | ASD | Multiplex, 2 siblings affected | Griswold et al., 2012 (22543975) |
| Copy number loss | P | 7851 | N/A | chr16:6926261-7102997 | 176,736 | UTR, exonic, intronic | exon 4, introns 3-4 (NM_018723) / introns 3-4, UTR, exons 1-2 (NM_001142334) | ASD | Simplex | Griswold et al., 2012 (22543975) |
| Copy number loss | P | 17855 | N/A | chr16:6926261-7102997 | 176,736 | UTR, exonic, intronic | exon 4, introns 3-4 (NM_018723) / introns 3-4, UTR, exons 1-2 (NM_001142334) | ASD | Simplex | Griswold et al., 2012 (22543975) |
| Copy number loss | M | 37350 | N/A | chr16:6926261-7102997 | 176,736 | UTR, exonic, intronic | exon 4, introns 3-4 (NM_018723) / introns 3-4, UTR, exons 1-2 (NM_001142334) | ASD | Unaffected sibling with CNV has a questionable autism diagnosis. | Griswold et al., 2012 (22543975) |
| Copy number loss | M | AU1393301 | N/A | chr16:6986112-7147772 | 16,166 | exonic, intronic | introns 1-2, UTR, exon 2 (NM_018723) / exon 2, introns 3-4 (NM_001142334) | ASD | N/A | Girirajan et al., 2013 (23375656) |
| Copy number loss | M | AU015903 | N/A | chr16:7109350-7211262 | 101,912 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Girirajan et al., 2013 (23375656) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Copy number loss** | M | 11464.p1 | N/A | chr16:6805493-6906683 | 101,190 | UTR, exonic, intronic | UTR exon 1, intron 1 (NM_001142334) / intron 3 (NM_018723) | ASD | N/A | Girirajan et al., 2013 (23375656) |
| **Copy number loss** | M | 13861.p1 | N/A | chr16:7566550-7652705 | 86,155 | exonic, intronic | exons 2-6, introns 1-6 (NM_145892) / exons 3-7, introns 2-7 (NM_001142334) / exons 5-9, introns 4-9 (NM_018723) | ASD | N/A | Girirajan et al., 2013 (23375656) |
| **Copy number loss** | M | 14073.p1 | N/A | chr16:7104795-7126957 | 22,162 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Girirajan et al., 2013 (23375656) |
| **Copy number loss** | M | 13414.p1 | N/A | chr16:6366866-6388848 | 21,982 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | ASD | N/A | Girirajan et al., 2013 (23375656) |
| **Copy number loss** | N/A | Case 2 | M | chr16:6458587-6477166 | 18,579 | intronic | intron 2 (NM_018723) | ASD, GDD, macroencephaly, (hypothyroid at birth, history of lactic acidosis) | N/A | Zhao 2013 (23822903) |
| **Copy number loss** | N/A | ASD0976 | M | chr16:6819005-6849571 | 30,567 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | ASD | N/A | Kushima 2018 et al., (30208311) |
| **Copy number loss** | N/A | ASD1001 | F | chr16:6819005-6849571 | 30,567 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | ASD | N/A | Kushima 2018 et al., (30208311) |
| **Copy number loss** | M | 1199_3 | M | chr16:6131347-6139749 | 8,403 | intronic | intron 1 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 14330_4440 | M | chr16:6884746-6963770 | 79,025 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 14162_2660 | M | chr16:6905072-6920592 | 15,521 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 16041_1571054001 | M | chr16:6593516-6633080 | 39,565 | intronic | intron 2 (NM_018723) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | M | 1948_301 | M | chr16:7213392-7238774 | 25,383 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Multiplex | Pinto et al., 2014 (24768552) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Copy number loss** | P | 20068_132 4001 | M | chr16:7106360-7130492 | 24,133 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | M | 3200_3 | M | chr16:6107421-6131347 | 23,927 | intronic | intron 1 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 3512_3 | M | chr16:7130492-7156382 | 25,891 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 3621_4 | F | chr16:6865181-7046258 | 181,078 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4182_1 | M | chr16:6869312-6927242 | 57,931 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4210_1 | M | chr16:6780569-7089819 | 309,251 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | Inherited | 4234_1 | M | chr16:7213392-7238774 | 25,383 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4305_1 | M | chr16:6926261-7102997 | 176,737 | exonic, intronic | intron 3, exon 4 (NM_018723) / intron 1, UTR, exon 2 (NM_001142334) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4418_1 | M | chr16:7126629-7199460 | 72,832 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Pinto et al., 2014 (24768552) |
| **Copy number loss** | M | 4461_1 | M | chr16:6862011-6966553 | 104,543 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4518_1 | M | chr16:6868286-6915483 | 47,198 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 4541_1 | M | chr16:6824823-6876984 | 52,162 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 5204_5 | M | chr16:6640396-6662435 | 22,040 | intronic | intron 2 (NM_018723) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| **Copy number loss** | P | 5297_3 | M | chr16:6356187-6369513 | 13,327 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | ASD | N/A | Pinto et al., 2014 (24768552) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | Inherited | 5355_3 | M | chr16:6873788-6879217 | 5,430 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| Copy number loss | M | 5424_3 | M | chr16:7213392-7238774 | 25,383 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| Copy number loss | M | 6408_3 | M | chr16:6850714-6892619 | 41,906 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| Copy number loss | M | 8642_201 | M | chr16:7062445-7082045 | 19,601 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | N/A | Pinto et al., 2014 (24768552) |
| Copy number loss | M | 8649_201 | M | chr16:6435694-6464810 | 29,117 | intronic | intron 2 (NM_018723) | ASD | N/A | Pinto et al., 2014 (24768552) |
| Copy number loss | Inherited | 20000_1010002 | F | chr16:7405611-7452970 | 47,360 | intronic | intron 4 (NM_145892) / intron 2 (NM_001142334) / intron 1 (NM_018723) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| Copy number loss | M | 284819 | F | chr16:7087824-7326867 | 239,040 | exonic | introns 3-4, exon 4 (NM_018723) / introns 1-2, exon 2 (NM_001142334) | ASD, moderate GDD | N/A | DECIPHER |
| Copy number loss | P | 290094 | N/A | chr16:6218528-6453034 | 234,510 | exonic | introns 1-2, exon 2 (NM_018723) | ASD, ID mild | N/A | DECIPHER |
| Copy number loss | de novo mosaic | 396533 | M | chr16:60001-16792499 | 16,730,000 | whole gene | whole gene | ASD, delayed speach and language development, ID, premature birth | N/A | DECIPHER |
| Copy number loss | N/A | 332117 | M | chr16:6819602-6869428 | 49,830 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_001142334) | ASD | N/A | DECIPHER |
| Copy number loss | N/A | 359240 | M | chr16:7070630-7382963 | 312,330 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) / UTR, exon 1 (NM_145892) | ASD | N/A | DECIPHER |
| Copy number loss | N/A | 331617 | M | chr16:6951071-7454813 | 503,740 | exonic | introns 3-4, exon 4 (NM_018723) / introns 1-2, exon 2 (NM_001142334) / UTR, exon 1, intron 1 (NM_145892) | ASD, overgrowth | N/A | DECIPHER |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Copy number loss** | Inherited | 257433 | M | chr16:6885122-7036209 | 151,090 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD, ID | Unaffected carrier parent | DECIPHER |
| **Copy number loss** | N/A | 385049 | M | chr16:5874625-8221258 | 2,350,000 | whole gene | whole gene | ASD, ID | N/A | DECIPHER |
| **Copy number loss** | N/A | 295208 | M | chr16:6707318-6895960 | 188,640 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_001142334) | ASD, hypertelorism | N/A | DECIPHER |
| **Copy number loss** | Inherited | 258954 | M | chr16:7376320-749386 | 117,670 | UTR, exonic, intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) / UTR, exon 1, intron 1 (NM_145892) | ASD, generalized tonic-clonic seizures, ID | N/A | DECIPHER |
| **Copy number loss** | Inherited | 256704 | M | chr16:69998726-7046219 | 47,490 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD, delayed speech and language development, ID | N/A | DECIPHER |
| **Copy number loss** | N/A | 249576 | N/A | chr16:3352432-6399846 | 3,050,000 | whole gene | whole gene | ASD, behavioural abnormality, morphological abnormalities | N/A | DECIPHER |
| **Copy number loss** | N/A | 380335 | M | chr16:6889408-7112742 | 223,340 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD with high cognitive abilities, tics | N/A | DECIPHER |
| **Copy number loss** | M | 276630 | M | chr16:6191542-6388677 | 197,140 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | ASD | N/A | DECIPHER |
| **Copy number loss** | N/A | 250901 | M | chr16:6554252-6673352 | 119,100 | intronic | intron 2 (NM_018723) | ASD, morphological abnormalities | N/A | DECIPHER |
| **Copy number loss** | M | 13122.p1 | N/A | chr16:6908075-7079700 | 171,625 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | N/A | Turner et al., 2016 (26749308) |
| **Copy number loss** | N/A | 12-4425-001 | N/A | chr16:7032320-7138711 | 106,392 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| **Copy number loss** | N/A | 12-4425-003 | N/A | chr16:7032320-7143879 | 111,560 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | N/A | 12-4425-004 | N/A | chr16:7032320-7138711 | 106,392 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | N/A | 12-4425-005 | N/A | chr16:7032320-7143879 | 111,560 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | N/A | 12-4425-006 | N/A | chr16:7032320-7143879 | 111,560 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | N/A | 12-4425-007 | N/A | chr16:7032320-7143879 | 111,560 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | N/A | 14-0260-003 | N/A | chr16:7101376-7131548 | 30,173 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | N/A | 7-0213-004 | F | chr16:6291019-6394434 | 103,416 | exonic, intronic | introns 1-2, exon 2(NM_018723) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | De Novo | AU077504 | F | chr16:6044487-6259852 | 215,365 | exonic, intronic | UTR, exon 1 , intron 1 (NM_018723) | ASD | sibship | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU3858302_HI10595 | M | chr16:6295074-6396948 | 101,874 | exonic, intronic | intron 1-2, exon 2 (NM_018723) | ASD | Trio | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU036204 | M | chr16:6337314-6614263 | 276,949 | exonic, intronic | intron 1-2, exon 2 (NM_018723) | ASD | Trio | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU036203 | M | chr16:6343042-6605213 | 262,171 | exonic, intronic | intron 1-2, exon 2 (NM_018723) | ASD | Trio | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU1393301 | M | chr16:6988411-7151333 | 162,922 | exonic, intronic | introns 1-2, UTR, exon 2 (NM_018723) / exon 2, introns 3-4  (NM_001142334) | ASD | Quad | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU1393303 | M | chr16:6988411-7151333 | 162,922 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | Quad | Leppa et al., 2016 (27569545) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | M | AU1393306 | M | chr16:6988411-7152865 | 164,454 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_001142334) | ASD | Quad | Leppa et al., 2016 (27569545) |
| Copy number loss | M | AU015903 | M | chr16:7109009-7211975 | 102,966 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Trio | Leppa et al., 2016 (27569545) |
| Copy number loss | N/A | nssv577542 | N/A | chr16:6295159-6377303 | 82,15 | exonic, intronic | introns 1-2, exon 2b(NM_018723) | ASD, GDD | N/A | ISCA database |
| Copy number loss | N/A | nssv577544 | N/A | chr16:6377244-7245430 | 868,19 | UTR, exonic, intronic | introns 2-4, exons 3-4 (NM_018723) / introns 1-2, UTR, exons 1-2 (NM_001142334) | ASD | N/A | ISCA database |
| Copy number loss | N/A | nssv582782 | N/A | chr16:6149327-6434200 | 284,87 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | ASD | N/A | ISCA database |
| Copy number loss | N/A | nssv1602798 | N/A | chr16:6972285-7004363 | 32,08 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | N/A | ISCA database |
| Copy number loss | N/A | nssv1602805 | N/A | chr16:6920984-6951130 | 30,15 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | N/A | ISCA database |
| Copy number loss | N/A | nssv1603768 | N/A | chr16:6972285-7026424 | 54,14 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD | N/A | ISCA database |
| Copy number loss | N/A | nssv1604541 | N/A | chr16:6889302-6964113 | 74,81 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | ASD, GDD | N/A | ISCA database |
| Copy number loss | N/A | nssv1605019 | N/A | chr16:7338230-7405379 | 67,15 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD, GDD, high-arched palate, muscular hypotonia | N/A | ISCA database |
| Copy number loss | N/A | 6046 | N/A | chr16:7102997-7126732 | 24 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | BIP | N/A | Noor et al., 2014 (24700553) |
| Copy number loss | N/A | 210 | N/A | chr16:7106360-7126732 | 20.373 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | BIP | N/A | Noor et al., 2014 (24700553) |
| Copy number loss | N/A | 6198-1 | N/A | chr16:6891681-6922307 | 30.626 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | BIP | N/A | Georgieva et al 2014 (25055870) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | M | N/A | N/A | chr16:6747759-6781695 | 33.937 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | BIP | N/A | Malhotra et al 2011 (22196331) |
| Copy number loss | M | 9025040001 | F | chr16:6294808–6394343 | 100 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | OCD | N/A | Grünblatt et al., 2017 (29179725) |
| Copy number loss | N/A | SCZ0613 | F | chr16:6691946-7089166 | 397,221 | UTR, exonic, intronic | introns 2-3, exon 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | SCZ | N/A | Kushima et al., 2018 (30208311) |
| Copy number loss | N/A | SCZ1786 | M | chr16:6682569-6780407 | 97,839 | exonic, intronic | introns 2-3, exon 3 (NM_018723) | SCZ | N/A | Kushima et al., 2018 (30208311) |
| Copy number loss | N/A | SCZ2138 | M | chr16:6833908-7038066 | 204,159 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | SCZ | N/A | Kushima et al., 2018 (30208311) |
| Copy number loss | N/A | SCZ2452 | M | chr16:6415715-6997801 | 582,087 | UTR, exonic, intronic | introns 2-3, exon 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | SCZ | N/A | Kushima et al., 2018 (30208311) |
| Copy number loss | N/A | SCZ2653 | F | chr16:6420350-6642206 | 221,857 | intronic | intron 2 (NM_018723) | SCZ | N/A | Kushima et al., 2018 (30208311) |
| Copy number loss | N/A | SCZ0839 | F | chr16:6714641-6792836 | 78,196 | intronic | intron 3 (NM_018723) | SCZ, reference and persecutory delusions, negative symptoms | N/A | Kushima et al., 2017 (27240532) |
| Copy number loss | N/A | case54 | M | chr16:6231252-6263633 | 32,382 | intronic | intron 1 (NM_018723) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | case85 | M | chr16:6814459-6865108 | 50,650 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | case86 | M | chr16:6826600-6835885 | 9,285 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | case87 | M | chr16:7033748-7138711 | 104,964 | exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, exon 2 (NM_01142334) | SCZ | N/A | Costain et al., 2013 (23813976) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Copy number loss | N/A | case62 | M | chr16:7052359-7196111 | 143,753 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_01142334) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | case202 | M | chr16:7107356-7130974 | 23,619 | intronic | intron 4 (NM_018723) / intron 2 (NM_01142334) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | case332 | M | chr16:6625422-6670858 | 45,437 | intronic | intron 2 (NM_018723) | SCZ | N/A | Costain et al., 2013 (23813976) |
| Copy number loss | N/A | N/A | N/A | chr16:6636120-6735137 | 99,017 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | deficit SCZ | N/A | Vrijenhoek et al., 2008 (18940311) |
| Copy number loss | M | N/A | N/A | chr16:6220199-6231993 | 11795 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | SCZ | N/A | Malhotra et al., 2011 (22196331) |
| Copy number loss | M | N/A | N/A | chr16:6531107-6547308 | 16202 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | SCZ | N/A | Malhotra et al., 2011 (22196331) |
| Copy number loss | N/A | 188856 | N/A | chr16:7052373-7197334 | 144,962 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_01142334) | SCZ | N/A | Zarrei et al., 2019 (31602316) |
| Copy number loss | M | N/A | F | chr16:6248324–6525864 | 277.540 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | Tourette-like syndrome, mild dysmorphic features | unaffected mother, brother with GAD and vocal tics (not genotyped) | Murgai et al., 2018 (30746397) |
| Copy number loss | N/A | TS1_0731 | F | chr16:6732433-7048405 | 315.973 | exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_001142334) | Tourette syndrome | N/A | Huang et al 2018 (28641109) |
| Copy number loss | N/A | TS2_1873 | M | chr16:6565650-6817375 | 251.726 | exonic, intronic | introns 2-3, exon 3 (NM_018723) | Tourette syndrome | N/A | Huang et al 2018 (28641109) |
| Copy number loss | N/A | 415959607 4_B | F | chr16:6859492-6876984 | 17.492 | intronic | intron 3 (NM_018723) / intron 1 (NM_001142334) | Tourette syndrome | N/A | Fernandez et al 2012 (22169095) |
| Copy number loss | M | Patient 3 | M | chr16:6218640-6294160 | 76,000 | intronic | intron 1 (NM_018723) | ASD, ADHD, oppositional defiant disorder and anxiety | mother with ADHD and dyslexia | Kamien et al., 2014 (24664471) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Copy number loss** | P | Patient 2 | M | chr16:6972277-7053842 | 82,000 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | ASD, anxiety, sensory issues | unaffected father but relatives with ASD and learning difficulties (no genotype data) | Kamien et al., 2014 (24664471) |
| **Copy number loss** | inherited | SCZ0076 | F | chr16:7089166-7159642 | 70,477 | exonic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_01142334) | SCZ, bizarre speech and behavior, irritability, anxiety | affected relatives (not details) | Kushima et al., 2018 (30208311) |
| **Copy number loss** | N/A | TS2_0758 | M | chr16:6813912-6901513 | 87.602 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | Tourette syndrome, OCD ADHD | N/A | Huang et al 2018 (28641109) |
| **Copy number loss** | N/A | TS2_0926 | F | chr16:6255416-6525983 | 270.568 | exonic, intronic | introns 1-2, exon 2 (NM_018723) | Tourette syndrome, OCD | N/A | Huang et al 2018 (28641109) |
| **Copy number loss** | N/A | TS2_1130 | F | chr16:6811163-7016549 | 205.387 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | Tourette syndrome, OCD ADHD | N/A | Huang et al 2018 (28641109) |
| **Copy number loss** | N/A | TS2_1641 | F | chr16:7091954-7232272 | 140.319 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, UTR, exon 2 (NM_01142334) | Tourette syndrome, OCD, ADHD | N/A | Huang et al 2018 (28641109) |
| **Copy number gain** | P | 130_089_3 | N/A | chr16:6813259-6901513 | 88,255 | UTR, exonic, intronic | intron 3 (NM_018723) / UTR exon 1, intron 1 (NM_01142334) | ADHD | unaffected father | Elia et al., 2010 (19546859) |
| **Copy number gain** | M | 130_214_3 | N/A | chr16:7283180-7330338 | 47,159 | intronic | intron 4 (NM_018723) / intron 2 (NM_01142334) | ADHD | mother meets criteria for adulthood ADHD (ASRS) | Elia et al., 2010 (19546859) |
| **Copy number gain** | N/A | S299 | N/A | chr16:6664674-6683075 | 18401 | intronic | intron 2 (NM_018723) | ADHD | N/A | Jarick et al., 2014 (23164820) |
| **Copy number gain** | P | 1912 | N/A | chr16:7757373-7776124 | 18,752 | UTR, exonic, intronic | introns 14-15, exons 15-16, UTR (NM_018723) / introns 12-13, exons 13-14, UTR (NM_001142334) / introns 11-12, exons 12-13, UTR (NM_145892) | ASD | N/A | Kanduri et al., 2016 (26052927) |
| **Copy number gain** | P | Patient 1 | F | chr16:6149356-6348681 | 199,000 | intronic | intron 1 (NM_018723) | ASD | father with mild learning difficulties | Kamien et al., 2014 (24664471) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Copy number gain | P | 3395_004 | M | chr16:6654691-6689092 | 34,402 | intronic | intron 2 (NM_018723) | ASD | N/A | Pinto et al., 2010 (20531469) |
| Copy number gain | P | 3097_004 | F | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | N/A | Pinto et al., 2010 (20531469) |
| Copy number gain | P | 14158_2580 | M | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | Simplex | Pinto et al., 2010 (20531469) |
| Copy number gain | M | 3095_003 | F | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | N/A | Pinto et al., 2010 (20531469) |
| Copy number gain | M | 14229_3630 | M | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | Simplex | Pinto et al., 2010 (20531469) |
| Copy number gain | M | 18100_302 | M | chr16:6657851-6689092 | 31,242 | intronic | intron 2 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| Copy number gain | M | 3095_3 | F | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| Copy number gain | P | 3097_4 | F | chr16:6657305-6689092 | 31,788 | intronic | intron 2 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| Copy number gain | P | 3395_4 | M | chr16:6654691-6689092 | 34,402 | intronic | intron 2 (NM_018723) | ASD | Multiplex | Pinto et al., 2014 (24768552) |
| Copy number gain | P | 4532_1 | M | chr16:7133248-7196046 | 62,799 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | Simplex | Pinto et al., 2014 (24768552) |
| Copy number gain | de novo | 5442_3 | M | chr16:7304904-7325683 | 20,780 | intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) | ASD | N/A | Pinto et al., 2014 (24768552) |
| Copy number gain | M | U-2015 | M | chr16:5942659-7000800 | 1,058 | UTR, exonic, intronic | UTR, introns 1-3, exons 1-3 (NM_018723) / UTR, intron 1 (NM_001142334) | ASD | N/A | Chen et al., 2017 (28931914) |
| Copy number gain | N/A | 1-0794-003 | M | chr16:6749974-7122769 | 372,796 | UTR, exonic, intronic | exon 4, intron 3-4 (NM_018723) / UTR, exon 1-2, intron 1-2 (NM_001142334) | ASD | N/A | Zarrei et al., 2019 (31602316) |
| Copy number gain | N/A | 1-0965-003 | M | chr16:6677327-6719038 | 41,712 | exonic, intronic | intron 2-3, exon 3 (NM_018723) | ASD | N/A | Zarrei et al., 2019 (31602316) |

| Copy number gain | P | AU353130 1_HI9803 | F | chr16:5783316-6887689 | 1,104 | UTR, exonic, intronic | introns 1-3, UTR, exons 1-3 (NM_018723) / UTR, exon 1-2, intron 1-2 (NM_001142334) | ASD | Trio | Leppa et al., 2016 (27569545) |
|---|---|---|---|---|---|---|---|---|---|---|
| Copy number gain | M | AU109130 6 | M | chr16:6104118-6257542 | 153,424 | intronic | intron 1 (NM_018723) | ASD | Quad | Leppa et al., 2016 (27569545) |
| Copy number gain | P | AU348230 3_HI9700 | M | chr16:6750250-7122374 | 372,124 | UTR, exonic, intronic | introns 3-4, exon 4 (NM_018723) / introns 1-2, exons 1-2, UTR (NM_001142334) | ASD | Trio | Leppa et al., 2016 (27569545) |
| Copy number gain | N/A | nssv160252 32 | N/A | chr16:6660795-6685821 | 25,03 | intronic | intron 2 (NM_018723) | ASD | N/A | ISCA database |
| Copy number gain | N/A | 385 | N/A | chr16:5803924-6425752 | 621.829 | UTR, exonic, intronic | introns 1-2, UTR, exons 1-2 (NM_018723) | BIP | N/A | Noor et al., 2014 (24700553) |
| Copy number gain | N/A | N/A | N/A | chr16:7194151-7416598 | 222,447 | UTR, exonic, intronic | intron 4 (NM_018723) / intron 2 (NM_001142334) / UTR, exon 1 (NM_145892) | BIP | N/A | Grozeva et al 2010 (20368508) |
| Copy number gain | N/A | N/A | N/A | chr16:6550400-6914358 | 363,958 | exonic, intronic | exon 3, introns 2-3 | BIP | N/A | Grozeva et al 2010 (20368508) |
| Copy number gain | M | 2058 | M | chr16:6641118-6688103 | 46985 | intronic | intron 2 (NM_018723) | SCZ | unaffected mother | Melhem et al., 2011 (21982423) |
| Copy number gain | N/A | 17615 | M | chr16:6642792-6688103 | 45,311 | intronic | intron 2 (NM_018723) | SCZ | unaffected sibling | Melhem et al., 2011 (21982423) |
| Copy number gain | P | 19800 | M | chr16:6642792-6688103 | 45,311 | intronic | intron 2 (NM_018723) | SCZ | unaffected father | Melhem et al., 2011 (21982423) |
| Copy number gain | P* | 20248 | M | chr16:6642792-6705571 | 62,779 | intronic | intron 2 (NM_018723) | SCZ | N/A | Melhem et al., 2011 (21982423) |
| Copy number gain | de novo | N/A | M | chr16:6948175-6978875 | 30,700 | intronic | intron 3 (NM_018723) / intron 1 (NM_01142334) | SCZ, learning disability | N/A | Xu et al., 2008 (18511947) |
| Copy number gain | N/A | lrr_baf_bat ch2.txt_49 3 1 | N/A | chr16:5536965-7209601 | 1672.636 | UTR, exonic, intronic | introns 1-4, exons 1-4, UTR (NM_018723) / introns 1-2, UTR, exons 1-2, UTR (NM_01142334) | Tourette syndrome | N/A | McGrath et al 2015 (25062598) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Copy number gain** | N/A | 1922040225_B | M | chr16:6624363-6681184 | 56.821 | intronic | intron 2 (NM_018723) | Tourette syndrome | N/A | Fernandez et al 2012 (22169095) |
| **Copy number gain** | N/A | TS2_1886 | F | chr16:7756087-7812131 | 56045 | UTR, exonic, intronic | intron 14-15, exon 15-16 (NM_018723) / introns 12-13, exons 13-14 (NM_01142334) / introns 11-12, exons 12-13 (NM_145892) | Tourette, ADHD | N/A | Huang et al 2018 (28641109) |
| **Copy number gain** | de novo | 395669 | M | chr16:60001-16792499 | 16,730,000 | whole gene | whole gene | ASD, aggressive behaviour, ID, morphological abnormalities | N/A | DECIPHER |
| **Copy number gain** | M | 22415 | M | chr16:6641118-6692980 | 51862 | intronic | intron 2 (NM_018723) | schizoaphective bipolar | unaffected mother | Melhem et al., 2011 (21982423) |

F, female; M, male ; N/A, data not available. ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; BIP, bipolar disorder; GDD, general developmental disorder; ID, intellectual disability; SCZ, schizophrenia.

**Supplementary Table 5.** CNVs in *RBFOX1* identified in studies including cases and controls, frequencies and burden analyses.

| Study (PMID) | Disorder | Total number of cases | Total number of controls | Cases with *RBFOX1* CN gain | Freq. RBFOX1 CN gain in cases | Controls with *RBFOX1* CN gain | Freq *RBFOX1* CN gain in controls | Cases with *RBFOX1* CN loss | Freq. *RBFOX1* CN loss in cases | Controls with *RBFOX1* CN loss | Freq. *RBFOX1* CN loss in controls | Ratio *RBFOX1* CNVs in cases vs controls | Burden Test all CNVs | Burden Test CN gain | Burden Test CN loss |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Elia et al., 2010 (19546859) | ADHD | 335 | 2026 | 2 | 0.60% | 0 | 0.00% | 1 | 0.30% | 0 | 0.00% | - | 0.004 | 0.151 | 0.019 |
| Lionel et al.. 2011 (21832240) | ADHD | 248 | 2357 | 0 | 0.00% | 0 | 0.00% | 1 | 0.40% | 0 | 0.00% | - | 0.096 | - | 0.096 |
| Jarik et al., 2014 (23164820) | ADHD | 489 | 1285 | 1 | 0.20% | 8 | 0.62% | 3 | 0.61% | 4 | 0.31% | 0.9 : 1 | 1 | 1 | 0.296 |
| **TOTAL** | **ADHD** | **1072** | **5668** | **3** | **0.28%** | **8** | **0.14%** | **5** | **0.47%** | **4** | **0.07%** | **3.5 : 1** | - | - | - |
| Prasad et al., 2012 (23275889) | ASD | 676 | 5139 | 0 | 0.00% | 0 | 0.00% | 3 | 0.44% | 0 | 0.00% | - | 0.001 | - | 0.002 |
| Bacchelli et al., 2020 (32081867) | ASD | 128 | 363 | 0 | 0.00% | 0 | 0.00% | 2 | 1.56% | 2 | 0.55% | 2.8 : 1 | 0.278 | - | 0.287 |
| Sebat et al., 2007 (17363630 ); Martin et al., 2007 (17503474) | ASD | 195 | 196 | 0 | 0.00% | 0 | 0.00% | 1 | 0.51% | 0 | 0.00% | - | 0.497 | - | 0.495 |
| Griswold et al., 2012 (22543975) | ASD | 813 | 592 | 0 | 0.00% | 0 | 0.00% | 4 | 0.49% | 0 | 0.00% | - | 0.108 | - | 0.109 |
| Girirajan et al., 2013 (23375656) | ASD | 2588 | 580 | 0 | 0.00% | 0 | 0.00% | 6 | 0.23% | 0 | 0.00% | - | 0.298 | - | 0.298 |
| Kushima et al., 2018 (30208311) | ASD | 1108 | 2095 | 0 | 0.00% | 0 | 0.00% | 2 | 0.18% | 3 | 0.14% | 1.3 : 1 | 0.552 | - | 0.568 |
| Kanduri et al., 2016 (26052927) | ASD | 80 | 269 | 1 | 1.25% | 0 | 0.00% | 0 | 0.00% | 2 | 0.74% | 1.7 : 1 | 0.542 | - | 1 |
| Chen et al., 2017 (28931914) | ASD | 335 | 1093 | 1 | 0.30% | 0 | 0.00% | 0 | 0.00% | 0 | 0.00% | - | 0.231 | 0.234 | - |
| **TOTAL** | **ASD** | **5923** | **10327** | **2** | **0.03%** | **0** | **0.00%** | **18** | **0.30%** | **7** | **0.07%** | **5.0 : 1** | - | - | - |

| Study | Disorder | Cases N | Controls N | | | | | | | | | Ratio | p | p | p |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Grozeva et al 2010 (20368508) | BIP | 1697 | 2806 | 2 | 0.12% | 0 | 0.00% | 0 | 0.00% | 0 | 0.00% | | 0.142 | 0.150 | - |
| **TOTAL** | **BIP** | **1697** | **2806** | **2** | **0.12%** | **0** | **0.00%** | **0** | **0.00%** | **0** | **0.00%** | **-** | - | - | - |
| Grünblatt et al., 2017 (29179725) | OCD | 121 | 124 | 0 | 0.00% | 0 | 0.00% | 1 | 0.83% | 0 | 0.00% | - | 0.503 | - | 0.498 |
| **TOTAL** | **OCD** | **121** | **124** | **0** | **0.00%** | **0** | **0.00%** | **1** | **0.83%** | **0** | **0.00%** | **-** | - | - | - |
| Kushima et al., 2018 (30208311) | SCZ | 2458 | 2095 | 0 | 0.00% | 0 | 0.00% | 6 | 0.24% | 3 | 0.14% | 1.7 : 1 | 0.340 | - | 0.333 |
| Kushima et al., 2017 (27240532) | SCZ | 1699 | 824 | 0 | 0.00% | 0 | 0.00% | 3 | 0.18% | 1 | 0.12% | 1.4 : 1 | 0.610 | - | 0.607 |
| Costain et al., 2013 (23813976) | SCZ | 454 | 416 | 0 | 0.00% | 0 | 0.00% | 7 | 1.54% | 5 | 1.20% | 1.3 : 1 | 0.444 | - | 0.452 |
| **TOTAL** | **SCZ** | **4611** | **3335** | **0** | **0.00%** | **0** | **0.00%** | **16** | **0.35%** | **9** | **0.27%** | **1.3 : 1** | - | - | - |
| Huang et al., 2018 (28641109) | Tourette syndrome | 2434 | 4093 | 1 | 0.04% | 1 | 0.02% | 6 | 0.25% | 7 | 0.17% | 1.5 : 1 | 0.304 | 0.601 | 0.344 |
| Fernandez et al., 2012 (22169095) | Tourette syndrome | 460 | 1131 | 1 | 0.22% | 3 | 0.27% | 1 | 0.22% | 2 | 0.18% | 1.0 : 1 | 1 | 1 | 0.642 |
| McGrath et al., 2015 (25062598) | Tourette syndrome | 1086 | 1789 | 1 | 0.09% | 0 | 0.00% | 0 | 0.00% | 0 | 0.00% | 1 | 0.370 | 0.387 | - |
| **TOTAL** | **Tourette syndrome** | **3980** | **7013** | **3** | **0.08%** | **4** | **0.06%** | **7** | **0.18%** | **9** | **0.13%** | **1.4 : 1** | - | - | - |
| **TOTAL** | **ALL** | **17404** | **29273** | **10** | **0.06%** | **12** | **0.04%** | **47** | **0.27%** | **29** | **0.10%** | **2.3 : 1** | - | - | - |

This table includes only 18 studies out of 34 (from Supplementary Table 4) for which information about CNVs in controls was available. ADHD, attention deficit hyperactivity disorder; ASD, autism spectrum disorder; BIP, bipolar disorder; CN, copy number; CNVs, copy number variants; Freq., frequency; OCD, obsessive-compulsive disorder; PMID, PubMed ID; SCZ, schizophrenia. Burden analysis performed using PLINK v.1.07. Underlined values correspond to p-values <0.05

**Supplementary Table 6.** Number of regulatory elements, H3K4me1 and H3K27ac peaks (data from ENCODE), detected in intronic CNVs in patients.

| Patient | H3K4me1 peaks | H3K27ac peaks |
|---|---|---|
| 107204L | 0 | 0 |
| 1199_3 | 1 | 1 |
| 13122.p1 | 25 | 25 |
| 14073.p1 | 1 | 0 |
| 14162_2660 | 2 | 4 |
| 14330_4440 | 10 | 9 |
| 16041_1571054001 | 14 | 2 |
| 1948_301 | 0 | 0 |
| 19764.3 | 3 | 5 |
| 20000_1010002 | 13 | 34 |
| 20068_1324001 | 2 | 0 |
| 210 | 1 | 0 |
| 250901 | 25 | 8 |
| 256704 | 140 | 141 |
| 257433 | 22 | 21 |
| 3200_3 | 4 | 4 |
| 3512_3 | 9 | E |
| 3621_4 | 22 | 22 |
| 4159596074_B | 0 | 0 |
| 4182_1 | 3 | 5 |
| 4234_1 | 0 | 0 |
| 4418_1 | 16 | 4 |
| 44307 | 1 | 1 |
| 4461_1 | 10 | 10 |
| 4518_1 | 3 | 4 |
| 4541_1 | 15 | 11 |
| 5204_5 | 1 | 1 |
| 5355_3 | 0 | 1 |
| 5424_3 | 0 | 0 |
| 6046 | 1 | 0 |
| 6198-1 | 3 | 4 |
| 6264_3 | 6 | 3 |
| 6317_5 | 3 | 5 |
| 6408_3 | 14 | 8 |
| 68257 | 3 | 5 |
| 8642_201 | 5 | 7 |
| 8649_201 | 0 | 1 |
| AB74 | 5 | 3 |
| AU015903 | 18 | 4 |

| | | |
|---|---|---|
| Case 2 | 1 | 0 |
| N/A_1 | 6 | 2 |
| N/A_2 | 5 | 23 |
| N/A_3 | 0 | 0 |
| Patient 2 | 11 | 11 |
| Patient 3 | 16 | 5 |
| S169 | 14 | 19 |
| S180 | 1 | 0 |
| SCZ0839 | 12 | 29 |
| SCZ2138 | 36 | 32 |
| case202 | 2 | 0 |
| case332 | 7 | 3 |
| case54 | 10 | 4 |
| case86 | 1 | 1 |
| nssv1602798 | 3 | 8 |
| nssv1602805 | 7 | 5 |
| nssv1603768 | 9 | 11 |
| nssv1604541 | 10 | 9 |
| 130_214_3 | 2 | 4 |
| 14158_2580 | 6 | 3 |
| 14229_3630 | 6 | 3 |
| 17615 | 7 | 4 |
| 18100_302 | 6 | 3 |
| 1922040225_B | 10 | 4 |
| 19800 | 7 | 4 |
| 20248 | 7 | 4 |
| 2058 | 7 | 4 |
| 22415 | 7 | 4 |
| 3095_003 | 6 | 3 |
| 3095_3 | 6 | 3 |
| 3097_004 | 6 | 3 |
| 3097_4 | 6 | 3 |
| 3395_004 | 6 | 3 |
| 3395_4 | 6 | 3 |
| 4532_1 | 15 | 4 |
| 5442_3 | 0 | 1 |
| AU1091306 | 27 | 16 |
| N/A | 5 | 5 |
| Patient_1 | 36 | 15 |
| S299 | 6 | 3 |
| nssv1602532 | 6 | 3 |

**Supplementary Table 7.** Altered expression of *RBOX1* in brain regions of individuals with SCZ or ASD

| Gene Symbol / Transcript symbol | Disorder | FC | *p*-value | FDR | Probe | Brain region (*post mortem*) | Study (PMID) or GEO ID | Sample, cases vs. controls |
|---|---|---|---|---|---|---|---|---|
| *RBFOX1* | ASD | -1.20 | 7.11E-03 | **0.068** | N/A | frontal and temporal cortex | Parikshak et al., 2016 (27919067) | 48 ASD vs. 49 control |
| *RBFOX1* | ASD | -1.41 | 2.20E-02 | N/A | ILMN_1814316 | temporal cortex | Schwede et al., 2018 (29859039) | 15 ASD vs. 16 control |
| *RBFOX1* | ASD | -1.40 | 3.12E-02 | N/A | ILMN_2359168 | temporal cortex | Schwede et al., 2018 (29859039) | 15 ASD vs. 16 control |
| *RBFOX1* | ASD | -1.39 | 1.59E-02 | N/A | ILMN_1731507 | temporal cortex | Schwede et al., 2018 (29859039) | 15 ASD vs. 16 control |
| *RBFOX1* | ASD | -1.32 | N/A | **0.043** | ILMN_1731507 | prefrontal and temporal cortex | Schwede et al., 2018 (29859039) | 15 ASD vs. 16 control |
| *RBFOX1* | ASD | -1.10 | 9.07E-04 | **0.025** | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 51 ASD vs. 936 control |
| *RBFOX1-002* | ASD | -1.14 | 8.69E-06 | **0.006** | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 51 ASD vs. 936 control |
| *RBFOX1-017* | ASD | 1.17 | 3.06E-02 | 0.369 | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 51 ASD vs. 936 control |
| *RBFOX1-016* | ASD | 1.35 | 2.26E-02 | 0.325 | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 51 ASD vs. 936 control |
| *RBFOX1-002* | SCZ | -1.02 | 1.55E-02 | 0.170 | N/A | frontal and temporal cortex | Gandal et al., 2018 (30545856) | 559 SCZ vs. 936 control |
| *RBFOX1* | SCZ | 1.10 | 3.47E-02 | 0.278 | 221217_s_at | prefrontal cortex | GSE21138 | 30 SCZ vs. 29 control |
| *RBFOX1* | SCZ | -1.26 | 4.89E-02 | 0.328 | 1553422_s_at | striatum | GSE53987 | 18 SCZ vs. 18 control |
| *RBFOX1* | SCZ | -1.17 | 4.60E-04 | **0.094** | 221217_s_at | prefrontal cortex | GSE53987 | 15 SCZ vs. 19 control |
| *RBFOX1* | SCZ | -1.18 | 2.38E-02 | 0.129 | 221217_s_at | hippocampus | GSE53987 | 15 SCZ vs. 18 control |
| *RBFOX1* | SCZ | -1.04 | 4.33E-02 | 0.327 | 7993083 | cerebellum | GSE35978 | 44 SCZ vs. 50 control |
| *RBFOX1* | SCZ | -1.04 | 4.33E-02 | 0.327 | 7993083 | cerebellum | GSE35974 | 44 SCZ vs. 50 control |

N/A, data not available; FC, fold change; FDR, False Discovery Rate; SCZ, schizophrenia; ASD, autism spectrum disorder. In bold, overcoming multiple testing corrections at 10% FDR.

**Supplementary Table 8.** Sample characteristics and behavioral performance for the Flanker/Go-nogo and face matching task (human imaging genetics sample)

| | Flanker/Go-Nogo (n=324) | | | Face matching (n=313) | | |
|---|---|---|---|---|---|---|
| | **C-carrier** | **T/T carrier** | **Test-statistic** | **C-carrier** | **T/T carrier** | **Test-statistic** |
| ***Demographics*** | | | | | | |
| **Age, mean +/- SD** | 32,9 ±9,7 | 33,0 ±10,1 | $t(322)$=0,100, $p$=0,920 | 33,7 ±9,8 | 33,5 ±10,04 | $t(311)$=-0,111, $p$=0,912 |
| **Sex (males/females)** | 119/134 | 36/35 | $X^2$=0,299, p=0,584 | 114/129 | 36/34 | $X^2$=0,444, p=0,505 |
| **Education (years), mean +/- SD** | 15,4 ±2,4 | 15,1 ±2,6 | $t(322)$=-0,960, $p$=0,338 | 15,5 ±2,5 | 15,3 ±2,7 | $t(311)$=-0,640, $p$=0,523 |
| **Site (Berlin/Mannheim/ Bonn)** | 72/80/101 | 16/24/31 | $X^2$=0,989, p=0,610 | 73/69/101 | 16/24/30 | $X^2$=1,643, $p$=0,440 |
| **Handedness (right/left/both)** | 230/16/6 | 62/8/1 | n/a* | 220/16/6 | 61/8/1 | n/a* |
| ***fMRI task performance*** | | | | | | |
| **Incongruent (% correct), mean ±SD** | 97,85 ±3,76 | 97,36 ±4,14 | F(1,317)=1,089, $p$=0,30[a,d] | | | |
| **Congruent (% correct), mean ±SD** | 99,42 ±2,44 | 98,42 ±4,68 | **F(1,317)=5,74, $p$=0,017** | | | |
| **Neutral (% correct), mean ±SD** | 99,07 ±3,12 | 98,73 ±4,51 | F(1,317)=0,54, $p$=0,46 | | | |
| **Nogo (% correct, no response), mean ±SD** | 91,65 ±6,4 | 91,55 ±5,75 | F(1,317)=0,02, $p$=0,90 | | | |
| **Incongruent, RT (ms), mean ±SD** | 632,78 ±117,07 | 623,60 ±96,71 | F(1,317)=0,62, $p$=0,41[a,b,c] | | | |
| **Congruent, RT (ms), mean ±SD** | 576,19 ±112,00 | 571,44 ±94,93 | F(1,317)=0,23, $p$=0,63[a,b,c] | | | |
| **Neutral, RT (ms), mean ±SD** | 608,33 ±113,65 | 597,40 ±94,94 | F(1,317)=0,87, $p$=0,35[a,b,c] | | | |
| **Faces (% correct)** | | | | 98,58 ±3,46 | 98,45 ±3,11 | F(1,306)=0,02, $p$=0,90 |
| **Forms (% correct)** | | | | 97,39 ±3,93 | 95,89 ±4,68 | **F(1,306)=7,04, $p$=0,008[b]** |

Type of statistical test for group comparisons: Normally distributed variables → independent t-tests or univariate ANOVAs including genotype and sex (between-subjects factors), and age and site (covariates of no interest), dichotomous variables → χ² test. Besides genotype effects univariate ANOVAs showed significant effects of [a]sex, [b]age, [c]site, [d]genotype x sex interaction. [#]statistical test not possible due to low number of "both" handedness. *Please note*: We controlled all second-level SPM12 analyses for the effects of sex, age, site, and behavioral covariates corresponding to the fMRI contrast of interest (see methods for details). Abbreviations: ms = milliseconds, RT = reaction time, SD = standard deviation.

**Supplementary Table 9.** Fear conditioning: Sociodemographic and psychological characteristics of panic disorder patients with C/C, C/T and T/T RBFOX1 rs6500744 SNP genotypes

| | C/C (n=15) | C/T (n=21) | T/T (n=11) | *F/Chi²* | *Post-hoc tests* |
|---|---|---|---|---|---|
| Age in years | 40.67±8.00 | 38.14±10.70 | 31.45±9.77 | 2.98 | |
| Female gender | 11 (73%) | 15 (71%) | 7 (64%) | 0.31 | |
| Years of education | | | | 1.05 | |
| ≤ 8 | 1 | 2 | 1 | | |
| 9 — 11 | 8 | 11 | 4 | | |
| ≥ 12 | 6 | 8 | 6 | | |
| Study Center | | | | 1.20 | |
| Center 1 | 4 | 7 | 4 | | |
| Center 2 | 1 | 1 | 0 | | |
| Center 3 | 3 | 5 | 2 | | |
| Center 4 | 7 | 8 | 5 | | |
| Digit span forward | 7.60±2.20 | 7.95±1.69 | 7.45±2.77 | 0.23 | |
| Digit span backward | 7.00±2.10 | 7.00±2.05 | 7.18±2.04 | 0.03 | |
| TMT-A | 25.60±5.17 | 27.50±9.99 | 21.54±4.12 | 2.22 | |
| TMT-B | 58.27±18.40 | 56.29±18.71 | 54.18±16.41 | 0.16 | |
| CGI | 5.53±0.74 | 5.43±0.60 | 5.36±0.51 | 0.25 | |
| SIGH-A | 25.67±6.02 | 23.33±4.48 | 24.09±5.67 | 0.86 | |
| MI alone | 2.90±0.91 | 2.34±1.22 | 2.45±0.91 | 1.29 | |
| PAS | 29.83±7.35 | 24.61±8.62 | 27.18±10.95 | 1.53 | |
| ASI | 37.00±7.31 | 27.95±9.64 | 29.45±10.02 | 4.63* | C/C>C/T |
| BDI-II | 23.67±9.08 | 14.48±4.91 | 12.91±7.85 | 9.62*** | C/C>C/T, C/C>T/T |

ASI: Anxiety Sensitivity Index; BDI-II: Beck Depression Inventory-II; BSI: Brief Symptom Inventory; CGI: Clinical Global Impression; MI: Mobility Inventory; PAS: Panic and Agoraphobia-Scale; SIGH-A: Structured Interview Guide for the Hamilton Anxiety Scale; TMT: Trail Making Test; *: $p < 0.05$; ***: $p < 0.001$

**Supplementary Table 10.** Behavioural performance in the fear conditioning task in panic disorder patients

| | C/C (n=19) | | T/C (n=34) | | T/T (n=12) | |
|---|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | Mean | SD |
| **Aversiveness US** | 8.00 | 1.15 | 7.91 | 1.74 | 7.08 | 2.99 |
| **Fam.: Valence CS+** | 3.16 | 0.83 | 3.09 | 1.08 | 3.00 | 0.95 |
| **Fam.: Arousal CS+** | 2.21 | 0.71 | 2.09 | 1.08 | 2.58 | 1.00 |
| **Fam.: Valence CS-** | 3.11 | 0.94 | 2.94 | 1.28 | 3.42 | 1.08 |
| **Fam.: Arousal CS-** | 2.21 | 0.85 | 2.03 | 1.14 | 2.25 | 1.06 |
| **Acq.: Valence CS+** | 3.05 | 1.08 | 2.94 | 1.18 | 2.92 | 0.90 |
| **Acq.: Arousal CS+** | 2.37 | 0.90 | 2.21 | 1.25 | 2.25 | 0.97 |
| **Acq.: Valence CS-** | 3.16 | 1.01 | 3.00 | 1.39 | 3.42 | 1.00 |
| **Acq.: Arousal CS-** | 2.16 | 1.12 | 1.97 | 1.14 | 2.17 | 1.03 |
| **Ext.: Valence CS+** | 3.53 | 0.77 | 2.94 | 1.15 | 3.25 | 0.87 |
| **Ext.: Arousal CS+** | 1.95 | 0.91 | 1.94 | 1.07 | 1.92 | 1.38 |
| **Ext.: Valence CS-** | 3.42 | 0.96 | 3.06 | 1.25 | 3.50 | 1.09 |
| **Ext.: Arousal CS-** | 2.16 | 1.07 | 1.91 | 1.19 | 1.92 | 1.24 |

Acq.: acquisition phase; CS: conditioned stimulus; Ext.: extinction phase; Fam.: familiarization phase; US: unconditioned stimulus.

**Supplementary Table 11.** Neural correlates of rs6500744 SNP genotype differences in fear conditioning and extinction

| Anatomical region | Cluster extension | BA | Coordinates | | | t-value | no. voxels |
|---|---|---|---|---|---|---|---|
| | | | x | y | z | | |
| *C/C >T/T in simple fear conditioning (Contrast: CS+ in late familiarization < CS+ in late acquisition)* | | | | | | | |
| Right MFG | Bilateral dmPFC, dACC, SFG, right IFG | 9,10,24,32,46 | 44 | 50 | 20 | 4.80 | 3931 |
| Right MTG | Right STG, insula, precentral gyrus, rolandic operculum, SMG | 21,22,39,41 | 34 | -16 | 42 | 4.34 | 6093 |
| Occipital visual cortex | | 17,18 | -36 | -88 | 12 | 4.29 | 791 |
| Left postcentral gyrus | Left STG, MTG, IPL | 3,39,41,42 | -50 | -18 | 52 | 4.18 | 2604 |
| Left Insula | Left thalamus, MTG, putamen | | 8 | -30 | -14 | 4.17 | 2219 |
| Right lingual gyrus | Right calcarine gyrus, fusiform gyrus | 17,18,19 | 18 | -84 | 4 | 4.14 | 1539 |
| Precuneus | PCC, paracentral lobule | 31 | -8 | -48 | 42 | 3.71 | 1150 |
| Left precentral gyrus | | 9 | -44 | 14 | 32 | 3.60 | 306 |
| Left Amygdala (ROI) | | | -26 | -2 | -14 | 2.83 | 20 |
| *C/C >T/T in differential fear conditioning (Contrast: CS+ in late acquisition > CS- in late acquisition)* | | | | | | | |
| PCC | Precuneus | 7,31 | -14 | -50 | 40 | 4.32 | 1436 |
| Right angular gyurs | Right STG, MTG | 22,40 | 48 | -58 | 40 | 3.90 | 852 |
| dmPFC | SMA | 8 | 10 | 24 | 48 | 3.74 | 298 |
| Left angular gyrus | Left IPL | 40 | -52 | -62 | 44 | 3.69 | 464 |
| Lingual gyrus | Calcarine gyrus | 18 | 10 | -78 | 0 | 3.49 | 478 |
| STG | | 22 | -48 | -42 | 14 | 3.27 | 185 |
| Cuneus | | 19 | 16 | -80 | 30 | 3.04 | 157 |
| *C/C >T/T in fear extinction (Contrast: CS+ in late acquisition > CS+ in late extinction)* | | | | | | | |
| dmPFC | dACC, SFG, | 6,8,9,32 | 4 | 66 | 14 | 4.47 | 3207 |
| Precentral gyrus | | 6 | -36 | -2 | 44 | 3.55 | 261 |
| Right IFG | MFG | 45 | 38 | 22 | 24 | 3.30 | 637 |

Participants are patients with panic disorder. Neuroimage Coordinates are listed in MNI space. ROI: region of interest analysis was performed within bilateral amygdala. Significance level: uncorrected $p < 0.005$, cluster with at least 141 voxels. BA: brodmann areas; dACC; Dorsal anterior cingulate cortex; dmPFC: dorsomedial prefrontal cortex; IFG: Inferior frontal gyrus; IPL: Inferior parietal lobule; MFG: Middle frontal gyrus; MTG: Middle temporal gyrus; PCC; Posterior cingulate cortex; SFG: Superior frontal gyrus; SMA: Supplementary motor cortex; STG: Superior temporal gyrus.

**Supplementary table 12.** Table/Figure BAT_1 (hrt1ages = anticipation phase; hrt1eges = exposure phase; hrt1pges = recovery phase; heart rate in bpm)

| | | | Mittelwert | Standardfehler des Mittelwerts | Standardabweichung | Gültige Anzahl |
|---|---|---|---|---|---|---|
| rs6500744_groups | CC | hrt1ages | 80,39 | 2,15 | 12,52 | 34 |
| | | hrt1eges | 81,45 | 2,21 | 12,86 | 34 |
| | | hrt1pges | 77,35 | 1,76 | 10,28 | 34 |
| | TC | hrt1ages | 78,09 | 1,62 | 11,89 | 54 |
| | | hrt1eges | 79,37 | 1,77 | 12,97 | 54 |
| | | hrt1pges | 76,32 | 1,50 | 11,05 | 54 |
| | TT | hrt1ages | 76,23 | 1,05 | 4,44 | 18 |
| | | hrt1eges | 77,44 | 1,23 | 5,24 | 18 |
| | | hrt1pges | 77,19 | 1,24 | 5,25 | 18 |
| | Gesamt | hrt1ages | 78,51 | 1,09 | 11,22 | 106 |
| | | hrt1eges | 79,71 | 1,16 | 11,97 | 106 |
| | | hrt1pges | 76,80 | ,97 | 9,98 | 106 |

**Supplementary table 13.** Increase of heart rate from last minute of anticipation phase to first minute of exposure phase (delta bpm)

| | | | Mittelwert | Standardfehler des Mittelwerts | Standardabweichung | Gültige Anzahl |
|---|---|---|---|---|---|---|
| rs6500744_groups | CC | hrt1em01_diff | 5,42 | 1,96 | 11,44 | 34 |
| | TC | hrt1em01_diff | 2,25 | 1,23 | 9,06 | 54 |
| | TT | hrt1em01_diff | -,66 | 1,37 | 5,82 | 18 |
| | Gesamt | hrt1em01_diff | 2,77 | ,93 | 9,62 | 106 |

**Supplementary Table 14**. Information about the custom-made KASP assay (LGC Genomics), used for SNP-specific genotyping.

| Assay ID | FAM Allele | HEX Allele | Aliquot ID | Thermal Cycling Protocol | Sequence |
|---|---|---|---|---|---|
| rs6500744-Rbfox1 | C | T | 1217845236 | 61-55°C *touchdown* | CACTTAATTTAGGTTTTCCAGTGTGGAGCATATCAGGAAGAGAAATGGC AGACATATTACTTGTCTCAAGCATTATCCATCATTAGCGTCTGTTCTCT TTCTCCGACACACACACACACACACACACACACACACACACACACACAC CCCCTTATATTTTTCCCCTAGTACCTTCCTCACTGATCATATTGTTCTG TTAACACCCACCACCAAAACTACATCCTGGGTCCTCTCCACACTGCAGC TCTCAAGACGATCGTTTACTCTATTATCAAGCTTACCATTTATTTTATT TTCAGG[C/T]GGTTGTATTCATTATAATGCCATTAGCATGAGAAGTGG GGTCCTGTCACCAGGGCAACCAGGAGCACAGAGTGCACATTCCCTCCTT AGCTGATGTCTTGTTCCCTCTTGGGTGGACTCCAACTGTGTCTTTGATT CATTCCATTAGGTGGTCACTGGGATGTAGATGCTTGACAGCCTCCTTGT CCTTAAGTTGTCTGTCTTTATATCCTTTACTGGCCAAGATTTCTCTCTC CAGAATTCATTAATTTATTACCTTTTATTTCTGTAAGCATAACACATAT AATATTTCCAAGCATAT |

**Supplementary Table 15**. *RBFOX1* rs6500744 effect on *RBFOX1* gene expression in post-mortem tissue samples: Sample characteristics.

| SNP Genotype | Sample Size (% male) | Cohort Breakdown[a] | | | | | Age Mean ± SEM | Age: Statistics |
|---|---|---|---|---|---|---|---|---|
| | | MTS | HBS | PIT | HRV | MIA | | |
| CC | 10 (50.0) | 2 | 2 | 4 | 1 | 1 | 61.5±6.9 | |
| CT | 15 (53.3) | 4 | 3 | 2 | 3 | 3 | 68.2±4.7 | $F_{(2,28)}$=0.43 $p$=0.65 n.s. |
| TT | 6 (83.3) | - | - | 1 | 3 | 2 | 62.8±5.8 | |

[a]*Cohorts represent biorepositories from the brain bank network. MTS = Mount Sinai Brain Bank; HBS = Human Brain and Spinal Fluid Resource Center; PIT = Brain Tissue Donation Program at the University of Pittsburgh; HRV = Harvard Brain Tissue Resource Center; MIA = University of Miami Brain Endowment Bank.*

# Article 4. Pleiotropic contribution of *rbfox1* to psychiatric and neurodevelopmental phenotypes in a zebrafish model

**Summary in Spanish: "Contribución pleiotrópica de *rbfox1* a fenotipos psiquiátricos en un modelo de pez cebra"**

Estudios recientes han señalado a *RBFOX1* como un gen altamente pleiotrópico que contribuye al desarrollo de diferentes trastornos psiquiátricos y del neurodesarrollo. De hecho, variantes comunes y raras en *RBFOX1* se han relacionado con trastornos mentales. En este estudio demostramos que en pez cebra *rbfox1* presenta una expresión pan-neuronal durante el desarrollo, lo que sugiere un papel importante de este gen en el neurodesarrollo, y una expresión restringida en el cerebro anterior, más concretamente en el telencéfalo, hipotálamo y tálamo, regiones que están implicadas en el procesamiento de información sensorial y en el comportamiento. Para investigar el efecto de la deficiencia de *rbfox1* en el comportamiento utilizamos la línea *rbfox1*^sa15940 genoanulada. Demostramos que los mutantes *rbfox1*^sa15940 presentan hiperactividad, ansiedad, una disminución del *freezing*, altos niveles de agresividad y un comportamiento social alterado. Repetimos estos experimentos con otra línea genoanulada, *rbfox1*^del19, con un trasfondo genético diferente y demostramos que los mutantes *rbfox1*^del19 tienen un comportamiento de tigmotaxis similar a los mutantes *rbfox1*^sa15940 pero no son agresivos, presentan alteraciones mayores en el comportamiento social y niveles menores de hiperactividad. En conjunto, estos resultados sugieren que una deficiencia de *rbfox1* produce cambios en el comportamiento asimilables a los que presentan pacientes con trastornos psiquiátricos, lo que valida estas dos líneas genoanuladas como modelos de trastornos psiquiátricos. Además, demostramos que el trasfondo genético puede modular los efectos de mutaciones en *rbfox1* en el comportamiento y realzamos el efecto pleiotrópico de *rbfox1* en trastornos psiquiátricos.

**Reference:**

Antón-Galindo E, Adel M, López-Blanch L, Norton WHJ, Fernàndez-Castillo N, Bru Cormand B. Pleiotropic contribution of *rbfox1* to psychiatric and neurodevelopmental phenotypes in a zebrafish model. To be submitted.

# Pleiotropic contribution of *rbfox1* to psychiatric and neurodevelopmental phenotypes in a zebrafish model

Ester Antón-Galindo[1,2,3,4], Maja Adel[1], Laura López-Blanch[5], William HJ Norton[5], Noèlia Fernàndez-Castillo[1,2,3,4]*, Bru Cormand[1,2,3,4]*,

[1]Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain

[2]Centro de Investigación Biomédica en Red de Enfermedades raras (CIBERER), Spain

[3]Institut de Biomedicina de la Universitat de Barcelona, Barcelona, Spain

[4]Institut de recerca Sant Joan de Déu; Espluges de Llobregat, Spain

[5]Department of Neuroscience, Psychology and Behaviour, College of Life Sciences, University of Leicester, Leicester, UK

6Universitat Pompeu Fabra (UPF), Barcelona, Spain.

7Department of Neuroscience, Psychology and Behaviour, College of Life Sciences, University of Leicester, Leicester, UK

* equally supervised this work

## ABSTRACT

Recent research highlighted *RBFOX1* as a highly pleiotropic gene contributing to several psychiatric and neurodevelopmental disorders. Indeed, both rare and common variants in *RBFOX1* have been associated with psychiatric conditions, but the mechanisms underlying the pleiotropic effects of *RBFOX1* are not yet understood. Here we found that in zebrafish, *rbfox1* presents a pan-neuronal expression during developmental stages, suggesting a major role of this gene in neurodevelopment, and a restricted expression in adult forebrain areas, including telencephalon, hypothalamus and thalamus, regions with an important role in receiving and processing sensory information and in directing behaviour. Then, to investigate the effect of *rbfox1* deficiency in behaviour, we used *rbfox1*[sa15940], an *rbfox1* knockout line. We found that *rbfox1*[sa15940] mutants present hyperactivity, anxiety-like behaviour, a decreased freezing behaviour, higher levels of aggression and an altered social behaviour. We repeated these behavioural tests in a different *rbfox1* knockout line, *rbfox1*[del19], with a different genetic background, and found that *rbfox1* deficiency affects behaviour differently. *rbfox1*[del19] mutants present a similar thigmotaxis behaviour, but stronger alterations in social behaviour and lower levels of hyperactivity than *rbfox1*[sa15940] fish. Contrary to *rbfox1*[sa15940], *rbfox1*[del19] mutants do not show an aggressive behaviour. Taken together, these results suggest that *rbfox1* deficiency leads to changes in behaviour in zebrafish that can be assimilated to phenotypical alterations present in patients with different psychiatric conditions, which validates these *rbfox1* KO zebrafish lines

as models for psychiatric disorders. Importantly, we show that the genetic background can modulate the effect of *rbfox1* knockout mutations on the behavioural phenotype and we highlight the pleiotropic effects of *rbfox1* on psychiatric disorders.

## INTRODUCTION

RNA Binding Fox-1 Homolog 1 (*RBFOX1*, also referred to as *A2BP1* or *FOX1*) encodes a RNA splicing factor, specifically expressed in brain, heart and muscle in humans (GTEX, www.gtexportal.com). This gene regulates the expression and splicing of large gene networks and plays an important function in neurodevelopment [1,2]. Rare genetic variations, including point mutations and copy number variants have been reported in *RBFOX1* in patients with neurodevelopmental disorders, such as autism spectrum disorder (ASD) [3–5], and *RBFOX1* haploinsufficiency results in a syndrome characterized by an impaired neurodevelopment [6,7]. In addition, transcriptomic analysis of autistic brains revealed decreased levels of *RBFOX1* and dysregulation of *RBFOX1*-dependent alternative splicing [8]. *RBFOX1* has not only been related to neurodevelopmental conditions, but increasing evidence points to both rare and common variants in this gene as contributors to several psychiatric and neurological disorders [5,9–11]. Indeed, *RBFOX1* was the second most pleiotropic locus in the recent Psychiatric Genomics Consortium cross-disorder genome-wide association studies (GWAS) meta-analysis [12]. All these data suggest a major role of *RBFOX1* in psychopathology, although the mechanisms underlying its pleiotropic effects are not yet understood.

*RBFOX1* has an orthologue gene in zebrafish, *rbfox1*, encoding a protein with an 84% identity to the human one [13]. Similar to the human orthologue, *rbfox1* is mainly expressed in brain, but also present in heart at early developmental stages [13]. However, to date, its expression at later stages has not been investigated nor its role in zebrafish neurodevelopment and behaviour. In the last years, zebrafish has become a powerful model to study psychiatric disorders, due to its high genetic similarity to human and its well-defined behavioural phenotypes, that can be easily assessed in the laboratory and that can be assimilated to human psychiatric phenotypes [14–16].

To date, genetic studies have pointed to a pleiotropic contribution of *RBFOX1* to several psychiatric conditions. Here, we aim to investigate the effect of the loss of *rbfox1* function in behaviour using a zebrafish model and to better understand the mechanisms underlying its pleiotropic effects on the onset of neurodevelopmental and psychiatric disorders.

## MATERIAL AND METHODS

### Zebrafish strains, care and maintenance

Adult zebrafish and larvae (*Danio rerio*) were maintained at 28.5°C in 14:10 light-dark cycles following standard protocols. All experimental procedures were approved by the Animal Welfare and Ethical Review board of the Generalitat de Catalunya. Behavioural experiments were performed with two different *rbfox1* mutant strains with different genetic background. *rbfox1*[sa15940], with Tubingen Long-fin (TL) background, is a transgenic line obtained from the ZIRC institute that contains an intronic point mutation at the -2 position of a 3' acceptor splicing site of *rbfox1* (A>T, Chr3:28068329, GRCz11). *rbfox1*[del19], with Tubingen (TU) background, was created by using CRISPR/Cas9 genetic engineering and causes a frameshift deletion of 19 bp in exon 2 of *rbfox1*. (Homozygous knockout fish (*rbfox1*[sa15940/sa15940] and *rbfox1*[del19/del19]), heterozygous (*rbfox1*[sa15940/ +] and *rbfox1*[del19/+]) and wild-type (WT, rbfox1[+/+]) fish were used for all behavioural experiments.

### Gene expression analysis using Real-Time quantitative PCR (RT-qPCR)

Total RNA was extracted from the whole brain of 5 TL WT, 5 *rbfox1*[sa15940/+] and two *rbfox1*[sa15940 /sa15940] adult zebrafish to perform RT-qPCR. Results were normalised to the expression levels of two housekeeping genes: the ribosomal protein L13a (*rpl13*) and the elongation factor 1a (*elf1a*). The relative expression of the genes and the fold change were calculated using the $2^{-\Delta\Delta CT}$ comparative method [17,18].

### In situ hybridization (ISH)

A specific mRNA probe targeting *rbfox1* (NCBI Reference Sequence: NM_001005596) was prepared and ISH experiments were performed in embryos, larvae and dissected adult brains of TL WT.

*Preparation of rbfox1 mRNA probe*

Total RNA was extracted from whole frozen adult WT zebrafish brains using TRIzol and complementary DNA (cDNA) was then synthesized using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific). A 565 bp *rbfox1* amplicon was generated by PCR (5'-CTCCAACATCCCCTTCAGGT-3' and 5'-TCGTCCGTAACTGTCACTGT-3' primers) and cloned into a plasmid using the StrataClone PCR cloning Kit (Agilent). The plasmids were collected and purified using GeneJET Plasmid Maxiprep Kit (Thermo Scientific) and the product was sequenced to check the orientation of the insert in the plasmid and the identity of the sequence. The selected plasmid containing the *rbfox1* insert was linearized with *Pst*I restriction enzyme and the *rbfox1*

DIG-antisense RNA probe was then generated by *in vitro* transcription. The product was DNase treated, cleaned using sodium acetate/ethanol precipitation, and the final *rbfox1* probe was stored at -20°C.

*Preparation of the samples for ISH*

Embryos were treated with 1-phenyl 2-thiourea (PTU) at 24 hours post fertilization (hpf) to prevent pigmentation. Embryos, larvae and dissected brains from adult fish were fixed overnight at 4°C in 4% paraformaldehyde (PFA) in phosphate-buffered saline (PBS). Specimens were then dehydrated with a gradient of methanol/PBS (25%, 50%, 75% and 100% methanol) before being stored for at least one hour and up to several months at -20°C.

*ISH protocol*

*First day of ISH*. Samples were rehydrated with a gradient of methanol/PBS (75%, 50%, 25% and 0% methanol) and then digested with proteinase K (10 µg/ml in PBS) at room temperature (25 minutes for 5 dpf embryos, 20 minutes for 4 dpf embryos, 15 minutes for 3 dpf embryos and 10 minutes for 2 dpf and 28 hpf embryos). Samples were then fixed in 4% PFA for 20 minutes and rinsed in phosphate-buffered saline + 0.1% Tween-20 (PBT). Samples were prehybridized at 68°C for at least 2 hours in 300 µl of HYB+ buffer (65% formamide, 5X saline-sodium citrate (SSC) buffer, 50 µg/ml heparin, 0.5 mg/ml torula RNA, 0.1% Tween-20, 9.2 mM citric acid, pH 6.0). HYB+ was then replaced with fresh HYB+ buffer containing the DIG-labelled probe (5 ng/µl) and incubated overnight at 68°C.

*Second day of ISH*. The HYB+/probe mix was removed and stored at -20°C for future use. Samples were washed with a gradient of HYB+/2X SSC (75%, 50%, 25% and 0% HYB) for 10 minutes each, and then twice with 0.05X SSC for 30 minutes each. For ISH on sections, adult brains were fixed for 20 min with 4% PFA and embedded in 3% agarose dissolved in water. Samples were sectioned at 100 µm using a vibratome and sections were collected in PBS. Specimens were blocked for 1 hour at room-temperature (RT) in blocking solution (2% normal goat serum, 2 mg/ml bovine serum albumin in PBT) and then incubated overnight with anti-DIG-AP antibody (1:4000 dilution in blocking solution).

*Third day of ISH*. Samples were washed several times in PBT and then three times for 10 minutes each in Xpho solution (100 mM Tris-HCl pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl and 0.1% Tween-20). Xpho solution was replaced with NBT/BCIP solution (225 µg/ml of NBT and 175µg/ml of BCIP in Xpho) and the specimens were incubated in the dark to develop the stain. Samples were monitored with a dissecting microscope every 30 minutes. The reaction was stopped by several washes in PBS and were fixed in 4% PFA for 20 minutes. Embryos were stored in 80% glycerol

and 20% PBT at 4°C, whereas sections were mounted on slides and covered with Mowiol solution. The NBT/BCIP signal was imaged using a GX microscope, a CMEX 5.0 camera and Image focus 4 software.

**Behavioural tests**

Behavioural analyses were performed on adult zebrafish (3-6 months-old) mixed groups of both sexes. All fish were genotyped, sized-matched and maintained in groups of 13 by genotype until the day of testing.  All the experiments were performed with homozygous knockout fish ($rbfox1^{sa15940/sa15940}$ and $rbfox1^{del19/del19}$), heterozygous ($rbfox1^{sa15940/+}$ and $rbfox1^{del19/+}$) and wild-type (TL rbfox1$^{+/+}$ and TU rbfox1$^{+/+}$) fish, and were completed between 9:00 and 18:00 and recorded using StreamPix 7 software (Norpix) and a digital camera. Fish were left for 30 minutes to habituate to the testing room before the experiment.

*Open field test*

The open field test was performed in a large circular open tank (43 cm of diameter) and the fish were recorded from above for 5 minutes. We used idtracker.ai and the *trajectorytools* module for Python to quantify the time spent in the centre of the tank, the time spent freezing, the distance swum and the velocity. We used 13 individuals per genotype.

*Shoaling test*

The shoaling experiment was performed following the protocol from Parker et al., 2013 [19]. We used idtracker.ai and the *trajectorytools* module for Python to measure the nearest neighbour distance (NDD), the inter-individual distance (IID), the distance swum and the velocity. We also virtually divided the tank into nine sections and calculated the cluster score across time [19]. We used two groups of five individuals per genotype.

*Visually-mediated social preference test (VMSP)*

The experiment was performed in a transparent tank composed of one central chamber (20 cm x 14 cm) surrounded by four identical chambers (10 cm x 7 cm). This test is divided in two steps as described in Carreño Gutierrez et al., 2019 [20]: social preference step and preference for social novelty step. During the first step (social preference), a first group of three unfamiliar WT fish were placed into one side of the side compartments. Then, the behaviour of a focal fish placed in the central area was recorded for five minutes. The time spent closer to the first group of strangers was compared to the time spent near the empty area diagonally opposite. In a second step (social novelty preference), a second group of three unfamiliar zebrafish were placed in the compartment diagonally opposite the first group. The focal fish was recorded for

five more minutes. We used a mixture of size-matched males and females as strangers since they can attract both male and female zebrafish [21]. We used idtracker.ai and the *trajectorytools* module for Python to measure the time spent in the different areas of the central compartment, the time spent freezing, the distance swum and the velocity. We used 13 individuals per genotype.

*Black and white test*

The black and white test was performed in a rectangular tank (24 cm x 12 cm) divided into two equal areas, a black area and a white area. Fish were placed in the centre of the tank and recorded for 5 minutes. The time spent in each area and the number of crossings between them were manually quantified. n=13/genotype.

*Aggression test*

Aggression was measured using the mirror-induced stimulation protocol [22]. A single fish was placed in the centre of tank with three white walls and a transparent wall, through which an external mirror can be seen, and was recorded for 5 minutes. The time spent in agonistic interaction with the mirror was manually quantified. We used 13 individuals per genotype.

## RESULTS

### *rbfox1* expression is pan-neuronal during development and restricted to specific brain areas during adulthood

*rbfox1* expression is widespread in brain during all developmental stages, as shown by ISH in 28 hpf to 5 dpf WT larvae (Figure 1A). In line with previous studies, we found that *rbfox1* is not only expressed in brain, but also in heart during development [13]. In adult brains, *rbfox1* expression is restricted to the forebrain, more precisely to ventral and dorsal telencephalic areas, paraventricular hypothalamus and thalamic nuclei (Figure 1B).
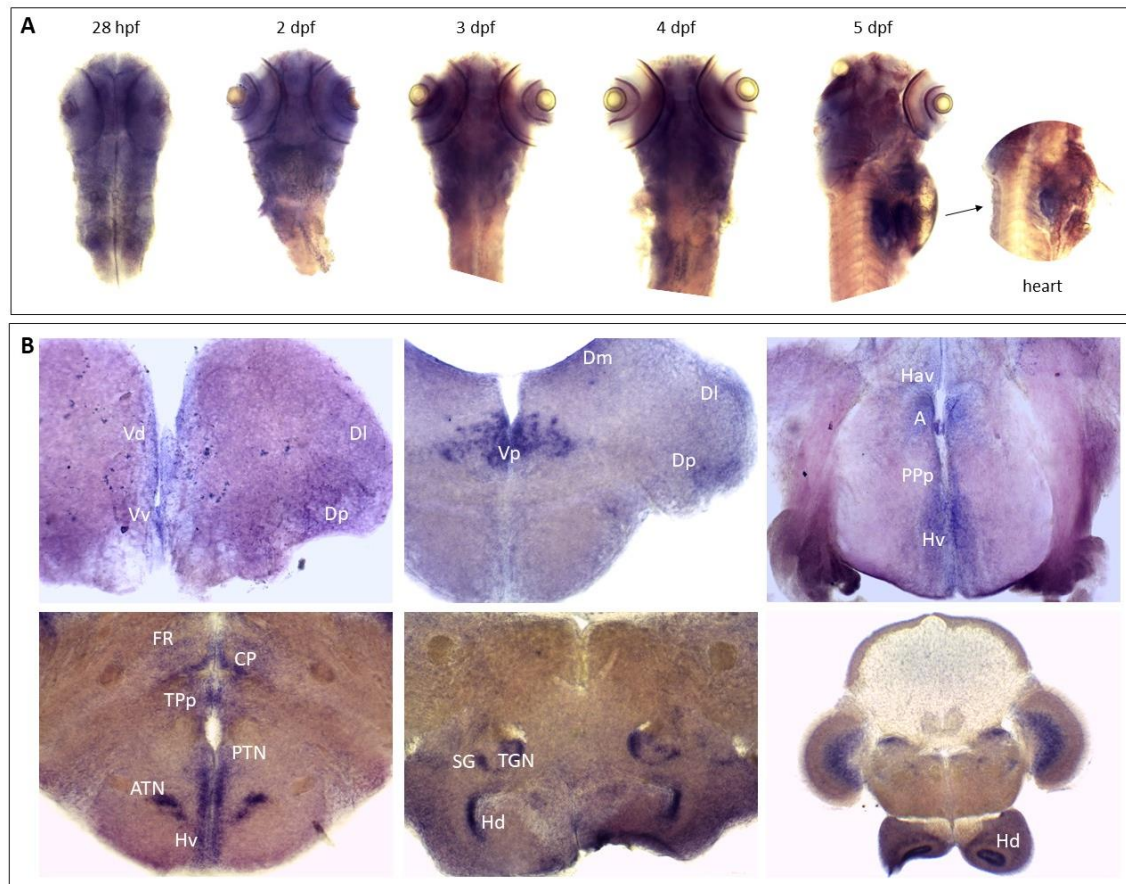
**Figure 1.** *rbfox1* **expression in zebrafish across development and in adult brain. A)** *rbfox1 in situ* hybridisation of whole-mount embryos from 28 hours post fertilization (hpf) until 5 days post fertilization (dpf). *rbfox1* is expressed pan-neuronally during larval stages, and in heart at 5 dpf. **B)** *rbfox1 in situ* hybridisation in adult brain cross-sections shows that *rbfox1* expression is restricted to the telencephalon, hypothalamus and thalamus nuclei. A, anterior thalamic nucleus; ATN, anterior tuberal nucleus; CP, central posterior thalamic nucleus; Dl, lateral zone of dorsal telencephalic area; Dm, medial zone of dorsal telencephalic area; Dp, posterior zone of dorsal telencephalic area; FR, fasciculus retroflexus; Hav, ventral habenular nucleus; Hd, dorsal zone of periventricular hypothalamus; Hv, ventral zone of periventricular hypothalamus; PPp, parvocellular preoptic nucleus, posterior part; PTN, posterior tuberal nucleus; SG, subglomerular nucleus; TGN, tertiary gustatory nucleus; TPp, periventricular nucleus of posteror tuberculum; Vd, dorsal zone of ventral telencephalic area; Vp, posterior zone of ventral telencephalic area; Vv, ventral zone of ventral telencephalic area.

### *rbfox1* is not expressed in the *rbfox1^sa1594/sa159400* zebrafish

*rbfox1^sa15940* (A>T, Chr3:28068329, GRCz11) is an intronic point mutation at the -2 position of the 3' acceptor splicing site affecting all but one of the *rbfox1* zebrafish isoforms (Figure 2). To check whether the splicing *rbfox1^sa15940* mutation in *rbfox1* triggers mRNA degradation by nonsense-mediated mRNA decay (NMD) [23] we analysed *rbfox1* mRNA levels by RT-qPCR. We observed

a significantly decreased level of *rbfox1* expression in *rbfox1^sa15940* mutants, both homozygous and heterozygous, compared to WT, using expression of *rpl13* and *elf1* as reference genes (Figure 2). These results suggest that NMD degradation of the truncated *rbfox1* transcript has occurred in mutants and that this mutant line could be used to examine the effect of loss of *rbfox1* function in zebrafish.
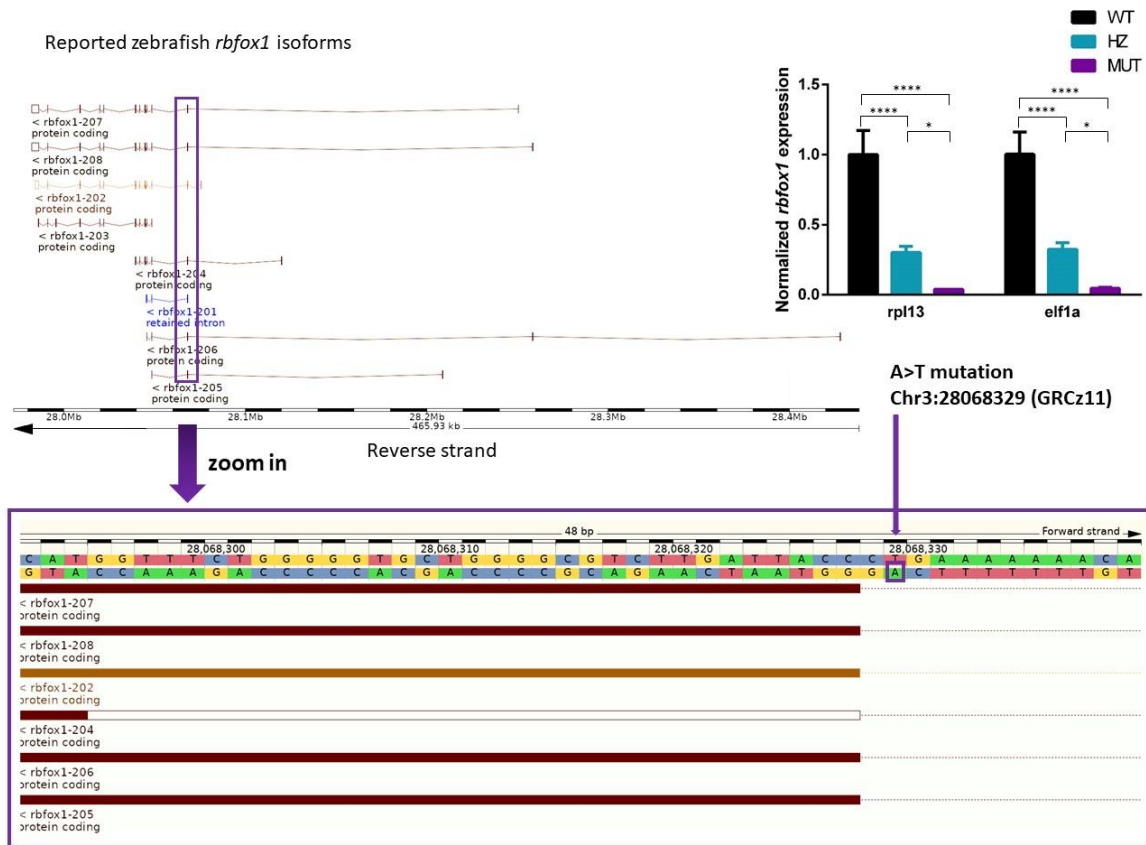


**Figure 2. sa15940 mutation in *rbfox1* gene: effects in *rbfox1* expression in adult brain.** Top left: *rbfox1* isoforms described in zebrafish (Ensembl database). Bottom: sa15940 is a point mutation (A>T, Chr3:28068329, GRCz11) situated in an intronic splicing region affecting all *rbfox1* isoforms described in zebrafish except for *rbfox1*-203. Top right: relative brain expression of *rbfox1* mRNA in adult fish. *rbfox1* expression is normalised to the average expression of *rbfox1* in wild-type (WT) fish and to two reference housekeeping genes: elongation factor 1a (*elf1a*) and ribosomal protein L13a (*rpl13*). *rbfox1^+/-sa15940* (HZ) fish present a reduction of around 70% *rbfox1* expression compared to WT (Mean HZ = 0.30 for rpl13 normalization / 0.32 for elf1a normalization; p < 0.0001, Two-way ANOVA followed by Sidak's multiple comparison test). *rbfox1^-/-sa15940* fish present a reduction of around 95% *rbfox1* expression compared to WT (Mean KO= 0.040 for rpl13 normalization / 0.045 for elf1a normalization; p < 0.0001, Two-way ANOVA followed by Sidak's multiple comparison test). n = 4 WT, 5 HZ, 2 KO. * p < 0.05; **** p < 0.0001. Mean ± SD.

**Loss of function of *rbfox1* produces behavioural alterations in *rbfox1$^{sa15940}$* zebrafish**

We performed a battery of five behavioural tests in WT TL, heterozygous (HZ) *rbfox1$^{sa15940/+}$* and homozygous (KO) *rbfox1$^{sa15940/ sa15940}$* adult fish, to investigate whether loss of *rbfox1* function affects behaviour.

For this mutant line, HZ and KO fish seem to be more anxious than WT TL fish, although not statistical differences were found in the time spent in the centre of the open field arena. All HZ and KO individuals spend less than 20% of the time in the centre and show a thigmotaxis behaviour, which does not occur in the WT group. In addition, HZ and KO fish spend less time freezing than WT fish and show hyperactivity, as they travel more distance and present a higher speed than WT TL individuals (Figure 3A and Supplementary Figure 1A).

In the VMSP test we did not observe differences in social preference between genotypes for this line. However, in the first step of the test, KO fish showed hyperactivity, reflected by more distance travelled and a higher speed than WT TL individuals (Figure 3B and Supplementary Figure 1B). Also, in the preference for social novelty step, a tendency to freeze more after the addition of the second group of strangers was seen in *rbfox1$^{sa15940}$* HZ and KO fish, being significant for the HZ group compared to WT TL (Figure 3C and Supplementary Figure 1B).

In the shoaling test, we only found differences in the mean interindividual distance (IID), that was higher in HZ and KO compared to WT TL fish (Figure 3D and Supplementary Figure 1D). No differences were found in the time spent in the white chamber of the black and white test, but KO fish cross more times the limit between areas, a sign of hyperactivity (Figure 3E). Finally, KO fish tend to be more aggressive than WT TL and are significantly more aggressive than HZ fish (Figure 3F).

Taken together, these results show behavioural alterations in *rbfox1*$^{sa15940}$ mutants, with TL genetic background, that present hyperactivity, an anxiety-like phenotype, alterations in social behaviour and a stronger aggressive behaviour.
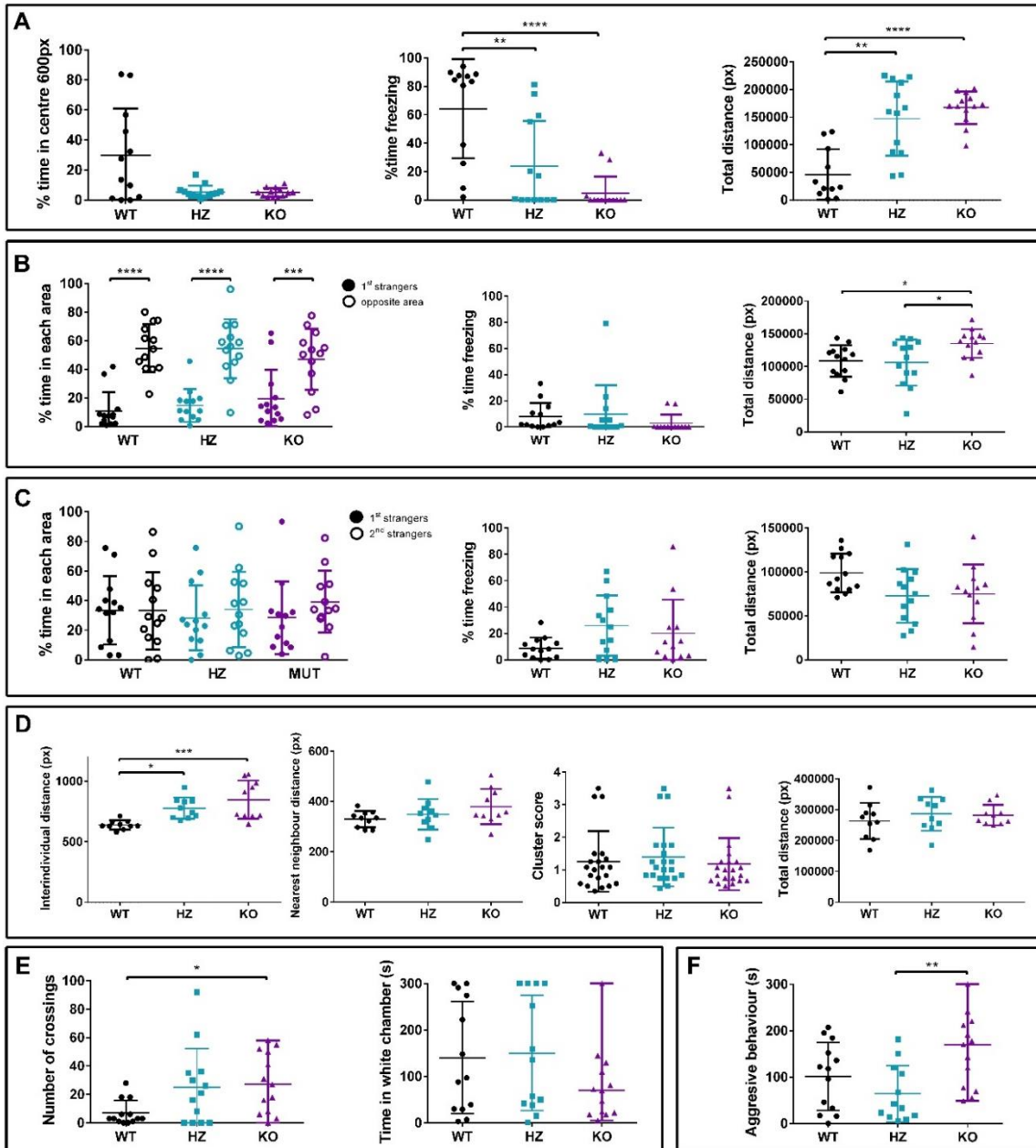
**Figure 3. Behavioural alterations observed in the *rbfox1*^sa15940 line. A) Open field test**. HZ and KO fish seem to be more anxious than WT fish as they spend less time in the centre of the arena, although differences observed are not statistically significant (WT vs. HZ, p = 0.38; WT vs. KO, p = 0.45). HZ and KO fish spend less time freezing than WT fish (WT vs. HZ, p = 0.0068; WT vs. KO, p = 0.0001) and travel more distance than WT individuals (WT vs. HZ, p = 0.0027; WT vs. KO, p = 0.0002). Kruskal-Wallis followed by Dunn's pot hoc tests. **B) Visually-mediated social preference test. Social preference step**. All the genotypes prefer to stay in the opposite corner rather than close to the group of stranger fish ($1^{st}$ strangers vs. Opposite area: WT, p < 0.0001; HZ, p < 0.0001; KO, p = 0.0005; Two-way ANOVA followed by Sidak's multiple comparison test). No differences were found in freezing behaviour, but KO fish travel more distance than WT and HZ individuals (HZ vs. KO, p = 0.0282; WT vs. KO, p = 0.0487; One-way ANOVA followed by Tuckey's multiple comparison test). **C) Visually-mediated social preference test. Preference for social novelty step**. All the genotypes show no preference for any group of stranger fish ($1^{st}$ strangers

vs. $2^{nd}$ strangers: WT, p > 0.99; HZ, p = 0.90; KO, p = 0.61; Two-way ANOVA followed by Sidak's multiple comparison test). No differences were found in freezing behaviour, although HZ and KO fish tend to freeze more than WT, or distance travelled. **D) Shoaling test**. Interindividual distance is higher between HZ and KO fish compared to WT (WT vs. HZ, p = 0.0194; WT vs. KO, p = 0.0005; One-way ANOVA followed by Tuckey's multiple comparison test). No differences were found between genotypes in nearest neighbour distance, cluster score or distance travelled. **E) Black and white test**. A higher number of crossings between areas was observed in KO fish, compared to WT (WT vs. KO, p = 0.0334; Kruskal-Wallis followed by Dunn's multiple comparisons test). No differences were observed in the time spent in each area. **F) Mirror test**. KO fish are more aggressive than WT and HZ fish, this difference being significant between HZ and KO fish (HZ vs. KO, p = 0.0083; Kruskal-Wallis followed by Dunn's multiple comparisons test). For all the experiments except for the shoaling test: n = 13 WT, 13 HZ and 13 KO. For the shoaling test: n = 2 groups of 5 individuals per genotype. * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001. Mean ± SD. KO, *rbfox1*[sa15940/sa15940] fish; HZ, *rbfox1* [sa15940/+] fish; WT, wild-type TU.

**The genetic background modulates the effect of loss of *rbfox1* function in behaviour**

We repeated the battery of behavioural tests in a different *rbfox1* mutant line, *rbfox1*[del19], to explore possible effects of the genetic background in the phenotypic expression of *rbfox1* deficiency. This mutant line was created by using CRISPR/Cas9 in zebrafish with a TU genetic background. In this line, the *rbfox1* mutation causes a frameshift deletion of 19 bp in exon 2 (affecting all rbfox1 zebrafish isoforms but rbfox1-203, see Figure 2) that produces NMD, as shown by qPCR (data not shown). We observed behavioural differences between *rbfox1*[del19] mutants and WT TU fish in all the tests performed, although the behavioural changes differed from the ones obtained for the *rbfox1*[sa19540] line.

We observed differences in the open field test between *rbfox1*[del19] and *rbfox1*[sa15940] lines. *rbfox1*[del19] mutants tend to spend less time in the centre than WT TU fish (being significant for HZ vs WT TU), but we did not find differences in locomotor activity between genotypes in the *rbfox1*[del19] line (nor in distance travelled or speed). In addition, we did not find differences in the freezing behaviour (Figure 4A and Supplementary Figure 2A).

Contrary to the *rbfox1*[sa19540] line, in the preference step of the VMSP test WT TU and HZ *rbfox1*[del19] fish show preference to stay close to stranger fish, whereas KO *rbfox1*[del19] fish show no social preference and spend significantly less time than WT TU fish near strangers and more in the opposite area (Figure 4B). In line with *rbfox1*[sa15940] results, KO *rbfox1*[del19] fish present hyperactivity, reflected by a higher speed than WT TU (Supplementary Figure 2B). In the social novelty preference step, we observed a similar behaviour in both *rbfox1*[del19] and *rbfox1*[sa15940]

lines: all the genotypes show no preference for a group of strangers. In the *rbfox1*[del19] line, KO fish present hyperactivity (Figure 4C and Supplementary Figure 2C).
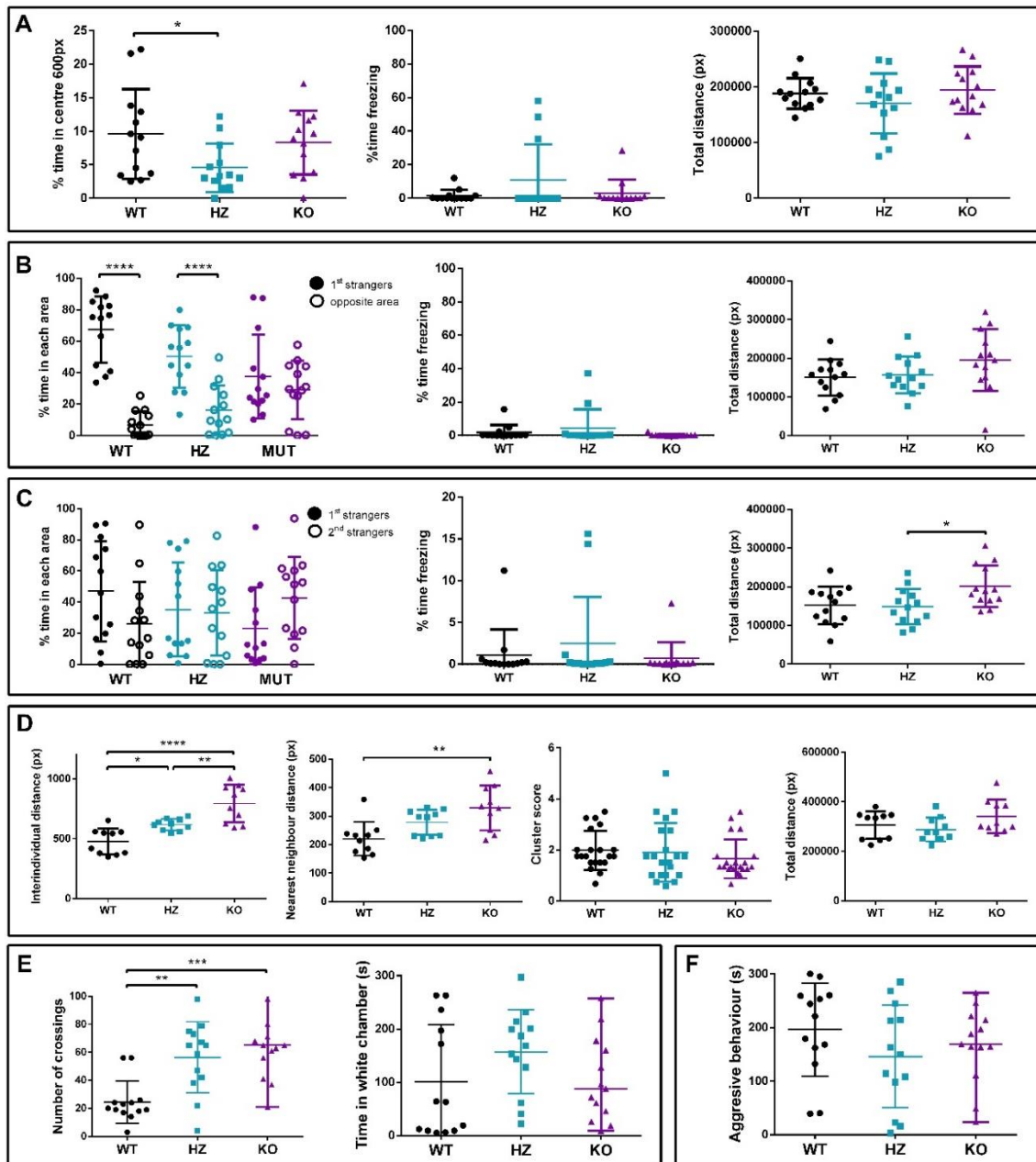


**Figure 4. Behavioural alterations observed in the *rbfox1*[del19] line. A) Open field test**. HZ and KO fish seem to be more anxious than WT fish as they spend less time in the centre of the arena, differences observed are statistically significant only for HZ fish (WT vs. HZ, p = 0.0467; One-way ANOVA followed by Tuckey's multiple comparison test). No differences were found in the time freezing, nor in the distance travelled. **B) Visually-mediated social preference test. Social preference step**. WT and HZ fish prefer to stay close to the group of stranger fish rather than in the opposite corner but KO fish show no preference for any of the areas (1st strangers vs. Opposite area: WT, p < 0.0001; p < 0.0001; KO, p = 0.6979; Two-way ANOVA followed by Sidak's multiple comparison test). No differences were found in freezing behaviour, nor in distance travelled. **C) Visually-mediated social preference test. Preference for social novelty step**. All the

genotypes show no preference for any group of stranger fish (1st strangers vs. 2nd strangers: WT, p = 0.18; HZ, p = 0.99; KO, p = 0.23; Two-way ANOVA followed by Sidak's multiple comparison test). No differences were found in freezing behaviour. KO fish travel more distance than both WT and HZ fish (WT vs. KO, p = 0.0467; HZ vs. KO, p = 0.0467; One-way ANOVA followed by Tuckey's multiple comparison test). **D) Shoaling test**. Interindividual distance is higher between HZ and KO fish compared to WT (WT vs. HZ, p = 0.0235; WT vs. KO, p < 0.0001; HZ vs. KO, p = 0.0047; One-way ANOVA followed by Tuckey's multiple comparison test). Nearest neighbour distance is higher between KO fish compared to WT (WT vs. KO, p < 0.0001; Kruskal-Wallis followed by Dunn's multiple comparisons test). No differences were found between genotypes in total distance travelled. **E) Black and white test**. A higher number of crossings between areas was observed in KO and HZ fish, compared to WT (WT vs. HZ, p = 0.0040; WT vs. KO, p = 0.0006; Kruskal-Wallis followed by Dunn's multiple comparisons test). No differences were observed in the time spent in each area.   **F) Mirror test**. No differences were observed in aggressive behaviour between genotypes. For all the experiments except for the shoaling test: n = 13 WT, 13 HZ and 13 KO. For the shoaling test: n = 2 groups of 5 individuals per genotype. * p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001. Mean ± SD. KO, *rbfox1*$^{del19/\,del19}$ fish;  HZ, *rbfox1*$^{del19/+}$ fish; WT, wild-type TU.

We found similar results in both *rbfox1* KO lines in the shoaling and black and white tests: mutant *rbfox1*$^{del19}$ fish present an impaired social behaviour and thigmotaxis (Figure 4D and Supplementary Figure 2D) and HZ and KO *rbfox1*$^{del19}$ performed a higher number of crossings between areas than WT (Figure 4E). Finally, contrary to KO *rbfox1*$^{sa15940}$ fish, KO *rbfox1*$^{del19}$ fish were not more aggressive than WT or HZ fish (Figure 4F).

In summary, both *rbfox1*$^{sa15940}$ and *rbfox1*$^{del19}$ mutants show thigmotaxis behaviour, an impaired social behaviour and hyperactivity. However, each *rbfox1* line present particularities: *rbfox1*$^{sa15940}$ mutants present alterations in freezing behaviour and aggression while *rbfox1*$^{del19}$ mutants present stronger social impairments. The behavioural differences reported between the two *rbfox1* mutant lines may be due to differences in the genetic background that modulate *rbfox1* effect on behaviour. Indeed, we can see that the genetic background is also influencing behaviour in the WT lines, as we observe strong differences in the freezing behaviour (Supplementary Figure 3).

## DISCUSSION

In this study we have investigated the role of *rbfox1* in neurodevelopmental and psychiatric disorders by studying the behavioural effects of loss of *rbfox1* function in zebrafish. This gene has previously been reported to be highly pleiotropic, contributing to several psychiatric

disorders [12,24]. In addition, we have validated zebrafish *rbfox1*[sa15940] and *rbfox1*[del19] KO lines as models of neurodevelopmental and psychiatric conditions.

First, *rbfox1* shows a restricted expression in brain and heart across developmental stages, and a pan-neuronal expression that suggests an important role of this gene during brain zebrafish development, in line with previous findings. Indeed, a study in human neural progenitor cells demonstrated that *RBFOX1* regulates splicing and expression of large gene networks implicated in neuronal development and maturation [25], and another study showed that *Rbfox1* controls synaptic transmission in the mouse brain [26,27]. Also, previous studies in mice have shown that specific *Rbfox1* deficiency in the central nervous system leads to impairments in neuronal migration, axon extension, dendritic arborisation and synapse network formation, suggesting that loss of *Rbfox1* function contributes to the pathophysiology of neurodevelopmental disorders [28–30]. Finally, several point mutations and CNVs in *RBFOX1* have been described in patients with neurodevelopmental disorders, such as ASD and attention-deficit hyperactivity disorder (ADHD) [4,5,9,31]. We therefore hypothesise that loss of *rbfox1* function may affect brain maturation in zebrafish and therefore lead to an impaired neuronal function and transmission during adulthood, with implications in the sensory response to the environment and in behaviour.

In addition, we found that *rbfox1* is specifically expressed in forebrain areas in adult WT animals, more precisely in dorsal and ventral telencephalon, thalamus and periventricular hypothalamus. Interestingly, these areas are involved in receiving and processing sensory information, stress, but also in directing behaviour, especially social behaviour and emotion [32–35]. Given the important role of *rbfox1* in controlling splicing and expression in neurons, *rbfox1* deficiency may induce an impaired neuronal function in these areas with an impact on sensory processing and behaviour in zebrafish.

Interestingly, both *rbfox1*[sa15940] and *rbfox1*[del19] KO lines present alterations in behaviour. *rbfox1*[sa15940] mutants present hyperactivity, thigmotaxis –an anxiety-like behaviour–, a decreased freezing behaviour, higher levels of aggression and an altered social behaviour. *rbfox1*[del19] mutants present a similar thigmotaxis behaviour, but stronger alterations in social behaviour and lower levels of hyperactivity than *rbfox1*[sa15940] fish. Contrary to *rbfox1*[sa15940], *rbfox1*[del19] mutants do not show an aggressive behaviour. All these behavioural phenotypes can be assimilated to phenotypical alterations observed in patients with psychiatric or neurodevelopmental conditions. For example, social impairment is a symptom of ASD, hyperactivity of ADHD, aggression is a phenotype associated to many psychiatric disorders and highly comorbid with ASD, and thigmotaxis is considered an anxiety-like behaviour in zebrafish.

These results point to a pleiotropic contribution of the *rbfox1* gene to neurodevelopmental and psychiatric disorders.

It is important to note the phenotypical differences observed between the two WT KO lines, that are probably due to the different genetic background. Behavioural differences between these two WT TL and TU strains have been previously reported, WT TL fish being considered as more anxious and sensitive to anxiogenic stimuli than WT TU fish [36]. Our results are in line with these reported phenotypes, as we found that WT TL presents a strong freezing behaviour, especially in the open field test, that is not present in WT TU fish (Supplementary Figure 3).

Given the differences observed between zebrafish lines, we hypothesise that loss of *rbfox1* function alters behaviour differently depending on the genetic background. Indeed, we found differences in behaviour between *rbfox1*<sup>sa15940</sup> and *rbfox1*<sup>del19</sup> KO lines. On one side, *rbfox1*<sup>sa15940</sup> is a hyperactive aggressive line that presents with anxiety-like behaviours and slight social impairments. On the other side, *rbfox1*<sup>del19</sup> fish are anxious but not aggressive, show only hyperactivity in one of the tests performed, but present stronger social impairments than *rbfox1*<sup>sa15940</sup> fish. These results suggest that the additive effects of variants in other genes contribute to the final phenotype, a genetic model that would be the rule in complex psychiatric disorders [37–40]. Our results show that the damaging effect of a loss-of-function mutation in *rbfox1* may be modulated by the genetic background and therefore lead to different phenotypes, which is in line with the different diagnosis of patients with rare CNVs or point mutations in the *RBFOX1* gene as well as the contribution of common variants to different psychiatric disorders [9,10,12,41].

To conclude, our results contribute to a better understanding of the involvement of *RBFOX1* in psychiatric disorders and point to a pleiotropic contribution of this gene that can be modulated by the individual genetic background. In addition, we have validated two new *rbfox1* KO zebrafish lines to be used as models of psychiatric disorders, in which further experiments can be performed to unravel the molecular mechanisms that link *RBFOX1* with psychiatric phenotypes.

## REFERENCES

1. Conboy, J.G. Developmental regulation of RNA processing by Rbfox proteins. *Wiley Interdiscip. Rev. RNA* 2017, *8*.

2. Bill, B.R.; Lowe, J.K.; DyBuncio, C.T.; Fogel, B.L. Orchestration of neurodevelopmental programs by RBFOX1: Implications for autism spectrum disorder. In *International Review of Neurobiology*; Academic Press Inc., 2013; Vol. 113, pp. 251–267.

3.  Bacchelli, E.; Cameli, C.; Viggiano, M.; Igliozzi, R.; Mancini, A.; Tancredi, R.; Battaglia, A.; Maestrini, E. An integrated analysis of rare CNV and exome variation in Autism Spectrum Disorder using the Infinium PsychArray. *Sci. Rep.* **2020**, *10*, 3198, doi:10.1038/s41598-020-59922-3.

4.  Griswold, A.J.; Dueker, N.D.; Van Booven, D.; Rantus, J.A.; Jaworski, J.M.; Slifer, S.H.; Schmidt, M.A.; Hulme, W.; Konidari, I.; Whitehead, P.L.; et al. Targeted massively parallel sequencing of autism spectrum disorder-associated genes in a case control cohort reveals rare loss-of-function risk variants. *Mol. Autism* **2015**, *6*, 1–11, doi:10.1186/s13229-015-0034-z.

5.  Zhao, W.W. Intragenic deletion of RBFOX1 associated with neurodevelopmental/ neuropsychiatric disorders and possibly other clinical presentations. *Mol. Cytogenet.* **2013**, *6*, 1–5, doi:10.1186/1755-8166-6-26.

6.  Bhalla, K.; Phillips, H.A.; Crawford, J.; McKenzie, O.L.D.; Mulley, J.C.; Eyre, H.; Gardner, A.E.; Kremmidiotis, G.; Callen, D.F. The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J. Hum. Genet.* **2004**, *49*, 308–311, doi:10.1007/s10038-004-0145-4.

7.  Martin, C.L.; Duvall, J.A.; Ilkin, Y.; Simon, J.S.; Arreaza, M.G.; Wilkes, K.; Alvarez-Retuerto, A.; Whichello, A.; Powell, C.M.; Rao, K.; et al. Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2007**, *144*, 869–876, doi:10.1002/ajmg.b.30530.

8.  Voineagu, I.; Wang, X.; Johnston, P.; Lowe, J.K.; Tian, Y.; Horvath, S.; Mill, J.; Cantor, R.M.; Blencowe, B.J.; Geschwind, D.H. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **2011**, *474*, 380–386, doi:10.1038/nature10110.

9.  Elia, J.; Gai, X.; Xie, H.M.; Perin, J.C.; Geiger, E.; Glessner, J.T.; D'Arcy, M.; Deberardinis, R.; Frackelton, E.; Kim, C.; et al. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry* **2010**, *15*, 637–646, doi:10.1038/mp.2009.57.

10. Kushima, I.; Aleksic, B.; Nakatochi, M.; Shimamura, T.; Shiino, T.; Yoshimi, A.; Kimura, H.; Takasaki, Y.; Wang, C.; Xing, J.; et al. High-resolution copy number variation analysis of schizophrenia in Japan. *Mol. Psychiatry* **2017**, *22*, 430–440, doi:10.1038/mp.2016.88.

11. Huang, A.Y.; Yu, D.; Davis, L.K.; Sul, J.H.; Tsetsos, F.; Ramensky, V.; Zelaya, I.; Ramos, E.M.; Osiecki, L.; Chen, J.A.; et al. Rare Copy Number Variants in NRXN1 and CNTN6 Increase Risk for Tourette Syndrome. *Neuron* **2017**, *94*, 1101-1111.e7, doi:10.1016/j.neuron.2017.06.010.

12. Lee, P.H.; Anttila, V.; Erneri; Won, H.; Feng, Y.-C.A.; Rosenthal, J.; Zhu, Z.; Tucker-Drob, E.M.; Nivard, M.G.; Grotzinger, A.D.; et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **2019**, *179*, 1469-1482.e11, doi:10.1016/j.cell.2019.11.020.

13. Frese, K.S.; Meder, B.; Keller, A.; Just, S.; Haas, J.; Vogel, B.; Fischer, S.; Backes, C.; Matzas, M.; Köhler, D.; et al. RNA splicing regulated by RBFOX1 is essential for cardiac function in zebrafish. *J. Cell Sci.* **2015**, *128*, 3030–3040, doi:10.1242/jcs.166850.

14. Fontana, B.D.; Mezzomo, N.J.; Kalueff, A. V.; Rosemberg, D.B. The developing utility of zebrafish models of neurological and neuropsychiatric disorders: A critical review. *Exp. Neurol.* **2018**, *299*, 157–171, doi:10.1016/j.expneurol.2017.10.004.

15. Norton, W. Towards developmental models of psychiatric disorders in zebrafish. *Front. Neural Circuits* **2013**, *7*, 1–

12, doi:10.3389/fncir.2013.00079.

16. Vaz, R.; Hofmeister, W.; Lindstrand, A. Zebrafish models of neurodevelopmental disorders: Limitations and benefits of current tools and techniques. *Int. J. Mol. Sci.* **2019**, *20*, doi:10.3390/ijms20061296.

17. Schmittgen, T.D.; Livak, K.J. Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* **2008**, *3*, 1101–1108, doi:10.1038/nprot.2008.73.

18. Livak, K.J.; Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods* **2001**, *25*, 402–408, doi:10.1006/meth.2001.1262.

19. Parker, M.O.; Brock, A.J.; Millington, M.E.; Brennan, C.H. Behavioural phenotyping of casper mutant and 1-Pheny-2-Thiourea treated adult zebrafish. *Zebrafish* **2013**, *10*, 466–471, doi:10.1089/zeb.2013.0878.

20. Carreño Gutiérrez, H.; Colanesi, S.; Cooper, B.; Reichmann, F.; Young, A.M.J.; Kelsh, R.N.; Norton, W.H.J. Endothelin neurotransmitter signalling controls zebrafish social behaviour. *Sci. Rep.* **2019**, *9*, doi:10.1038/s41598-019-39907-7.

21. Ruhl, N.; McRobert, S.P.; Currie, W.J.S. Shoaling preferences and the effects of sex ratio on spawning and aggression in small laboratory populations of zebrafish (Danio rerio). *Lab Anim. (NY).* **2009**, *38*, 264–269, doi:10.1038/laban0809-264.

22. Norton, W.H.J.; Stumpenhorst, K.; Faus-Kessler, T.; Folchert, A.; Rohner, N.; Harris, M.P.; Callebert, J.; Bally-Cuif, L. Modulation of fgfr1a signaling in zebrafish reveals a genetic basis for the aggression-boldness syndrome. *J. Neurosci.* **2011**, *31*, 13796–13807, doi:10.1523/JNEUROSCI.2892-11.2011.

23. Karousis, E.D.; Nasif, S.; Mühlemann, O. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley Interdiscip. Rev. RNA* **2016**, *7*, 661–682, doi:10.1002/wrna.1357.

24. Fernàndez-Castillo, N.; Gan, G.; van Donkelaar, M.M.J.; Vaht, M.; Weber, H.; Retz, W.; Meyer-Lindenberg, A.; Franke, B.; Harro, J.; Reif, A.; et al. RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur. Neuropsychopharmacol.* **2020**, *30*, 44–55, doi:10.1016/j.euroneuro.2017.11.012.

25. Fogel, B.L.; Wexler, E.; Wahnich, A.; Friedrich, T.; Vijayendran, C.; Gao, F.; Parikshak, N.; Konopka, G.; Geschwind, D.H. RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum. Mol. Genet.* **2012**, *21*, 4171–4186, doi:10.1093/hmg/dds240.

26. Vuong, C.K.; Wei, W.; Lee, J.A.; Lin, C.H.; Damianov, A.; de la Torre-Ubieta, L.; Halabi, R.; Otis, K.O.; Martin, K.C.; O'Dell, T.J.; et al. Rbfox1 Regulates Synaptic Transmission through the Inhibitory Neuron-Specific vSNARE Vamp1. *Neuron* **2018**, *98*, 127-141.e7, doi:10.1016/j.neuron.2018.03.008.

27. Gehman, L.T.; Stoilov, P.; Maguire, J.; Damianov, A.; Lin, C.H.; Shiue, L.; Ares, M.; Mody, I.; Black, D.L. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat. Genet.* **2011**, *43*, 706–711, doi:10.1038/ng.841.

28. Hamada, N.; Ito, H.; Iwamoto, I.; Morishita, R.; Tabata, H.; Nagata, K.I. Role of the cytoplasmic isoform of RBFOX1/A2BP1 in establishing the architecture of the developing cerebral cortex. *Mol. Autism* **2015**, *6*, 1–13, doi:10.1186/s13229-015-0049-5.

29. Hamada, N.; Ito, H.; Nishijo, T.; Iwamoto, I.; Morishita, R.; Tabata, H.; Momiyama, T.; Nagata, K.I. Essential role of the nuclear isoform of RBFOX1, a candidate gene for autism spectrum disorders, in the brain development. *Sci.*

*Rep.* **2016**, *6*, 1–19, doi:10.1038/srep30805.

30. Lee, J.A.; Damianov, A.; Lin, C.H.; Fontes, M.; Parikshak, N.N.; Anderson, E.S.; Geschwind, D.H.; Black, D.L.; Martin, K.C. Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. *Neuron* **2016**, *89*, 113–128, doi:10.1016/j.neuron.2015.11.025.

31. Pinto, D.; Delaby, E.; Merico, D.; Barbosa, M.; Merikangas, A.; Klei, L.; Thiruvahindrapuram, B.; Xu, X.; Ziman, R.; Wang, Z.; et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* **2014**, *94*, 677–694, doi:10.1016/j.ajhg.2014.03.018.

32. Teles, M.C.; Cardoso, S.D.; Oliveira, R.F. Social plasticity relies on different neuroplasticity mechanisms across the brain social decision-making network in zebrafish. *Front. Behav. Neurosci.* **2016**, *10*, doi:10.3389/fnbeh.2016.00016.

33. Stednitz, S.J.; McDermott, E.M.; Ncube, D.; Tallafuss, A.; Eisen, J.S.; Washbourne, P. Forebrain Control of Behaviorally Driven Social Orienting in Zebrafish. *Curr. Biol.* **2018**, *28*, 2445-2451.e3, doi:10.1016/j.cub.2018.06.016.

34. Mueller, T. What is the thalamus in zebrafish? *Front. Neurosci.* **2012**, *6*, 1–14, doi:10.3389/fnins.2012.00064.

35. Qin, C.; Li, J.; Tang, K. The paraventricular nucleus of the hypothalamus: Development, function, and human diseases. *Endocrinology* **2018**, *159*, 3458–3472, doi:10.1210/en.2018-00453.

36. Kalueff, A. V.; Stewart, A.M.; Gerlai, R. Zebrafish as an emerging model for studying complex brain disorders. *Trends Pharmacol. Sci.* **2014**, *35*, 63–75, doi:10.1016/j.tips.2013.12.002.

37. Niemi, M.E.K.; Martin, H.C.; Rice, D.L.; Gallone, G.; Gordon, S.; Kelemen, M.; McAloney, K.; McRae, J.; Radford, E.J.; Yu, S.; et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* **2018**, *562*, 268–271, doi:10.1038/s41586-018-0566-4.

38. Middeldorp, C.M.; Wray, N.R. The value of polygenic analyses in psychiatry. *World Psychiatry* **2018**, *17*, 26–28, doi:DOI:10.1002/wps.20480.

39. Boyle, E.A.; Li, Y.I.; Pritchard, J.K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **2017**, *169*, 1177–1186, doi:10.1016/j.cell.2017.05.038.

40. Liu, X.; Li, Y.I.; Pritchard, J.K. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell* **2019**, *177*, 1022-1034.e6, doi:10.1016/j.cell.2019.04.014.

41. Prasad, A.; Merico, D.; Thiruvahindrapuram, B.; Wei, J.; Lionel, A.C.; Sato, D.; Rickaby, J.; Lu, C.; Szatmari, P.; Roberts, W.; et al. A Discovery resource of rare copy number variations in individuals with autism spectrum disorder. *G3 Genes, Genomes, Genet.* **2012**, *2*, 1665–1685, doi:10.1534/g3.112.004689.

## SUPPLEMENTARY FIGURES
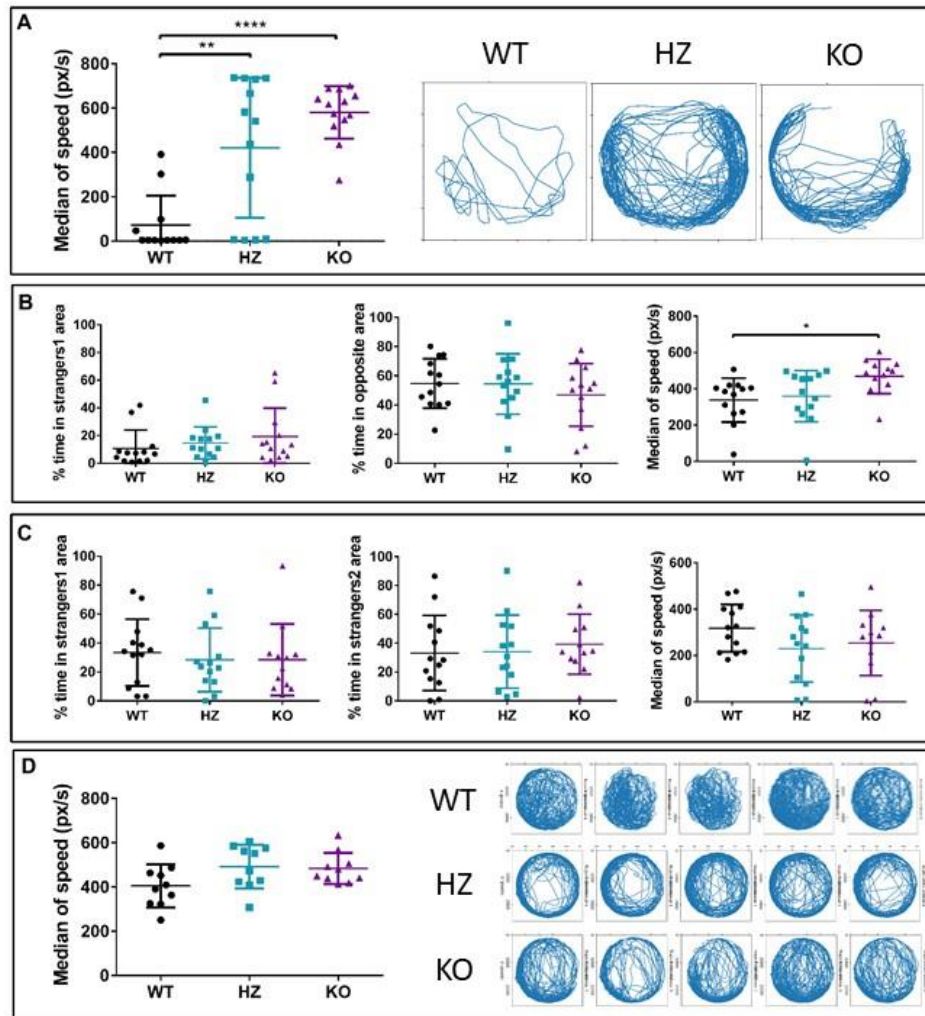


**Supplementary Figure 1**. **Behavioural alterations observed in *rbfox1*<sup>sa15940</sup>**. **A) Open field.** MUT and HZ fish swim at a higher speed than WT (WT vs. HZ, p = 0.0026; WT vs. MUT, p =0.0002; Kruskal-Wallis followed by Dunn's multiple comparisons test). Also, HZ and MUT fish show thigmotaxis behavior, as shown in these trajectories examples of different individuals. **B) Visually-mediated social preference test. Social preference step**. No differences were found between genotypes in the time spent near the 1$^{st}$ strangers or in the opposite corner. MUT fish swim at a higher speed than WT (WT vs. MUT, p =0.0130; Kruskal-Wallis followed by Dunn's multiple comparisons test). **C) Visually-mediated social preference test. Preference for social novelty step**. No differences were found between genotypes in the time spent near the 1$^{st}$ strangers or the 2$^{nd}$ strangers, nor in the velocity. **D) Shoaling test**. No differences were found between genotypes in the velocity, although there is a tendency in mutant fish to travel at a higher speed. Also, HZ and MUT fish present thighmotaxis behaviour, as shown in these trajectories examples of different individuals.  For all the experiments except for the shoaling test: n = 13 WT, 13 HZ and 13 MUT. For the shoaling test: n = 2 groups of 5 individuals per genotype. * p < 0.05; ** p < 0.01; **** p < 0.0001. Mean ± SD. KO, *rbfox1*$^{sa15940/sa15940}$ fish; HZ, heterozygous *rbfox1* $^{sa15940/+}$ fish; WT, wild-type TU.
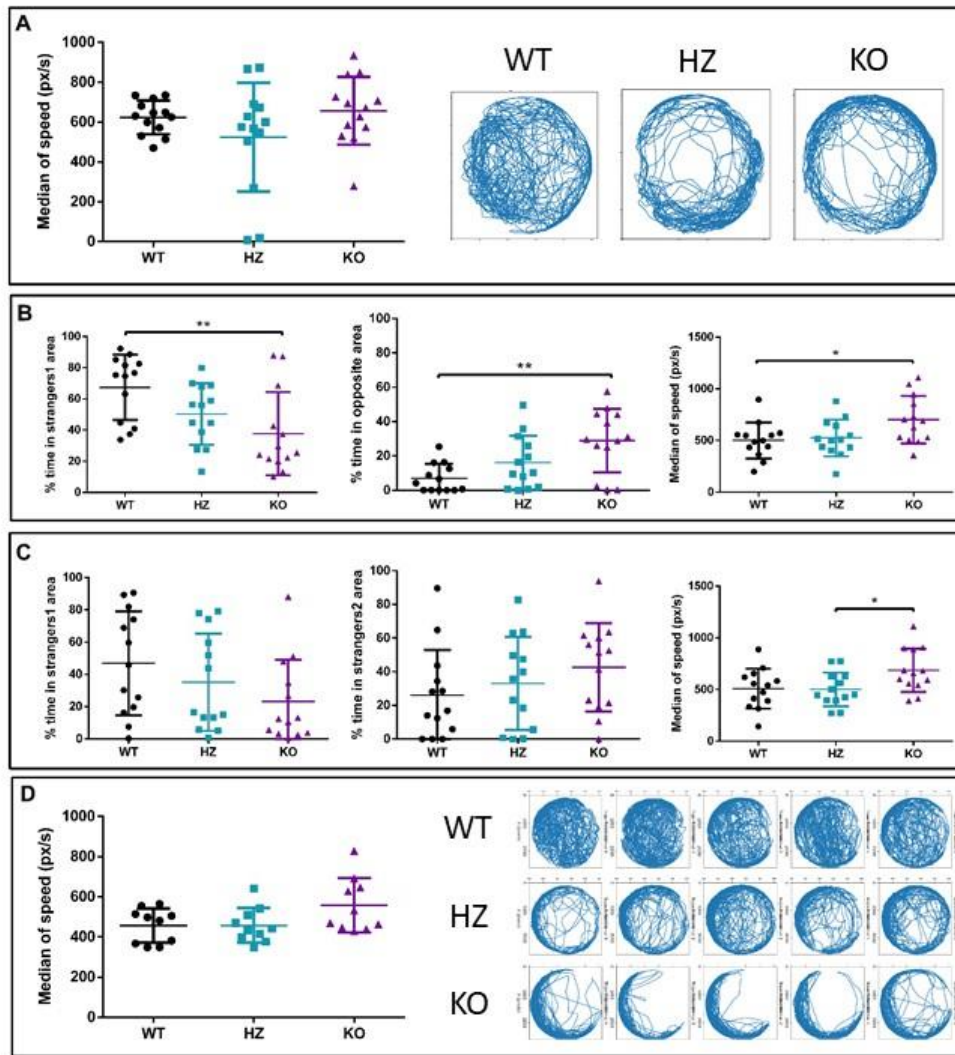
**Supplementary Figure 2**. **Behavioural alterations observed in *rbfox1*del19. A) Open field.** No differences were observed in speed between genotypes. However, HZ and MUT fish show thigmotaxis behavior, as shown in these trajectories examples of different individuals. **B) Visually-mediated social preference test. Social preference step**. MUT fish spend less time close to the 1st stranger fish and more time in the opposite area compared to WT animals (WT vs. MUT, p =0.0057; Kruskal-Wallis followed by Dunn's multiple comparisons test). Also, MUT fish swim at a higher speed than WT (WT vs. MUT, p =0.0339; Kruskal-Wallis followed by Dunn's multiple comparisons test). **C) Visually-mediated social preference test. Preference for social novelty step**. No significant differences were found between genotypes in the time spent near the 1st strangers or in the opposite corner, but MUT swim at a higher speed than HZ fish (HZ vs. MUT, p =0.0456; One-way ANOVA followed by Tuckey's multiple comparisons test). **D) Shoaling test**. No differences were found between genotypes in the velocity, although there is a tendency in mutant fish to travel at a higher speed. Also, HZ and MUT fish present thighmotaxis behaviour, as shown in these trajectories examples of different individuals.  For all the experiments except for the shoaling test: n = 13 WT, 13 HZ and 13 MUT. For the shoaling test: n = 2 groups of 5 individuals per genotype. * p < 0.05; ** p < 0.01. Mean ± SD. KO, *rbfox1*del19/ del19 fish;  HZ, heterozygous *rbfox1*del19/+ fish; WT, wild-type TU.
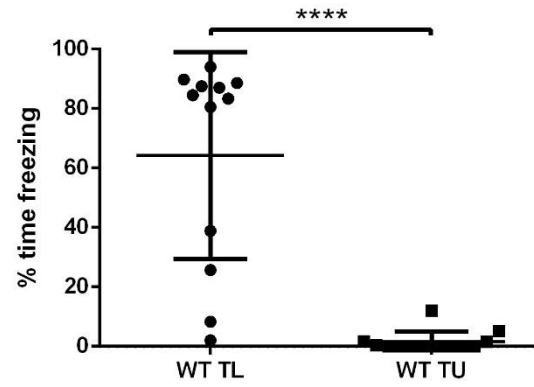
**Supplementary Figure 3. Comparison of the time freezing during the open-field test between the two wild-type lines used as controls in the behavioural experiments**. WT, wild-type; TL, Tubingen Long-fin line; TU, Tubingen line. **** $p < 0.0001$, Mann-Whitney U test. Mean ± SD.

# CHAPTER 3.

## Exploring the contribution of the *BEX/TCEAL* gene family

## to ASD and other psychiatric disorders

# Article 5. Characterization of an eutherian gene cluster generated after transposon domestication identifies *Bex3* as relevant for advanced neurological functions

**Summary in Spanish: "Caracterización de un grupo de genes euterios generado tras la domesticación de un transposón identifica a *Bex3* como gen relevante en funciones neurológicas avanzadas"**

**Introducción**: Una de las fuentes menos frecuentes de genes restringidos filogenéticamente es la domesticación molecular de elementos transponibles en un genoma hospedador. Pese a que estos eventos están a veces implicados en importantes cambios macroevolutivos, su origen y función no se han descrito en profundidad. **Resultados**: En este estudio identificamos eventos de domesticación que implican elementos transponibles no identificados previamente. Entre ellos, encontramos una domesticación molecular relevante que llevó a la formación de una familia génica en mamíferos placentarios: el grupo de genes *Bex/Tceal*. Estos genes, que actúan como proteínas centrales en varias vías de señalización, se han asociado con rasgos neurológicos en pacientes con microdeleciones en el cromosoma X. Además, los genes Bex/Tceal se expresan de forma diferencial en pacientes con trastorno del espectro autista y esquizofrenia. Por último, dos líneas murinas con mutaciones diferentes en el gen *Bex3* presentan alteraciones morfológicas y fisiopatológicas en cerebro: una reducción en el número de interneuronas y un desequilibrio electrofisiológico en hipocampo, alteraciones que se relacionan con un fenotipo de comportamiento alterado. **Conclusiones**: Reportamos el origen de un grupo de genes mediante domesticación de transposones y duplicación genética en mamíferos placentarios, un proceso evolutivo que transformó una secuencia no funcional en nuevos componentes del genoma euterio. Estos genes se integraron en vías preexistentes implicadas en el desarrollo, mantenimiento y función del sistema nervioso central en euterios. Además, demostramos que uno de los miembros de esta familia, el gen *Bex3*, está implicado en funciones cerebrales importantes en mamíferos placentarios y probablemente en trastornos neurológicos en humanos.

## Reference:

Navas-Pérez E, Vicente-García C, Mirra S, Burguera D, Fernàndez-Castillo N, Ferrán JL, López-Mayorga M, Alaiz-Noya M, Suárez-Pereira I, Antón-Galindo E, Ulloa F, Herrera-Úbeda C, Cuscó P, Falcón-Moya R, Rodríguez-Moreno A, D'Aniello S, Cormand B, Marfany G, Soriano E, Carrión ÁM, Carvajal JJ, Garcia-Fernàndez J. Characterization of an eutherian gene cluster generated after transposon domestication identifies Bex3 as relevant for advanced neurological functions. Genome Biol. 2020 Oct 26;21(1):267.

# Genome Biology

**RESEARCH**                                                                **Open Access**

# Characterization of an eutherian gene cluster generated after transposon domestication identifies *Bex3* as relevant for advanced neurological functions

Enrique Navas-Pérez[1†], Cristina Vicente-García[2†], Serena Mirra[1,3,4,5†], Demian Burguera[1,6], Noèlia Fernàndez-Castillo[1,4,7], José Luis Ferrán[8], Macarena López-Mayorga[2], Marta Alaiz-Noya[9,10], Irene Suárez-Pereira[9,11], Ester Antón-Galindo[1], Fausto Ulloa[3,5], Carlos Herrera-Úbeda[1], Pol Cuscó[12,13], Rafael Falcón-Moya[9], Antonio Rodríguez-Moreno[9], Salvatore D'Aniello[14], Bru Cormand[1,4,7], Gemma Marfany[1,4,7], Eduardo Soriano[3,5,15], Ángel M. Carrión[9], Jaime J. Carvajal[2*] and Jordi Garcia-Fernàndez[1*]

\* Correspondence: j.carvajal@csic.es;
jordigarcia@ub.edu
†Enrique Navas-Pérez, Cristina
Vicente-García and Serena Mirra
contributed equally to this work.
²Centro Andaluz de Biología del
Desarrollo, CSIC-UPO-JA,
Universidad Pablo de Olavide,
41013 Sevilla, Spain
¹Department of Genetics,
Microbiology and Statistics, Faculty
of Biology, and Institut de
Biomedicina (IBUB), University of
Barcelona, 08028 Barcelona, Spain
Full list of author information is
available at the end of the article

## Abstract

**Background:** One of the most unusual sources of phylogenetically restricted genes is the molecular domestication of transposable elements into a host genome as functional genes. Although these kinds of events are sometimes at the core of key macroevolutionary changes, their origin and organismal function are generally poorly understood.

**Results:** Here, we identify several previously unreported transposable element domestication events in the human and mouse genomes. Among them, we find a remarkable molecular domestication that gave rise to a multigenic family in placental mammals, the *Bex/Tceal* gene cluster. These genes, which act as hub proteins within diverse signaling pathways, have been associated with neurological features of human patients carrying genomic microdeletions in chromosome X. The *Bex/Tceal* genes display neural-enriched patterns and are differentially expressed in human neurological disorders, such as autism and schizophrenia. Two different murine alleles of the cluster member *Bex3* display morphological and physiopathological brain modifications, such as reduced interneuron number and hippocampal electrophysiological imbalance, alterations that translate into distinct behavioral phenotypes.

(Continued on next page)

(Continued from previous page)

**Conclusions:** We provide an in-depth understanding of the emergence of a gene cluster that originated by transposon domestication and gene duplication at the origin of placental mammals, an evolutionary process that transformed a non-functional transposon sequence into novel components of the eutherian genome. These genes were integrated into existing signaling pathways involved in the development, maintenance, and function of the CNS in eutherians. At least one of its members, *Bex3*, is relevant for higher brain functions in placental mammals and may be involved in human neurological disorders.

**Keywords:** Genetic novelty, Transposon domestication, *Bex3*, *Tceal*, Placental mammals, Gene cluster, Neurodevelopmental disorders, mTOR, Autism spectrum disorder

## Background

Newly evolved genes in a given lineage showing no homologs in other taxa are known as "orphan" or "taxonomically restricted" genes [1]. One of the most striking sources for the birth of lineage-restricted genes is the molecular domestication of transposable element (TE) proteins into novel coding genes [2], which are sometimes involved in the appearance of clade-specific traits and even true evolutionary novelties [3]. Despite their evolutionary relevance, a systematic search for domesticated transposons taking advantage of current and improved genomic annotations was lacking in human and mouse. We have identified now several domestication cases in these species. Among them, we highlight here a previously unreported event that took place at the origin of eutherian mammals, which gave rise to a multigenic family known as *Bex/Tceal* and shaped a cluster of 14 genes on the X chromosome of the placental ancestor. While most are scarcely studied, several reports have linked some of these genes to processes such as cancer proliferation [4–9], cellular reprogramming [10] and differentiation [11], or cell cycle modulation [12, 13]. To investigate the function of this gene family at the organism level, we generated and phenotyped mutant mice for one of its members, *Bex3*. Mutants showed molecular, cellular and anatomical alterations in the brain, as well as important neurological and behavioral alterations typical of neurodevelopmental defects. Altogether, we describe the evolutionary pathways of a TE-derived gene family that integrated into complex molecular routes, while underscoring its impact on neural development and its neuropsychiatric significance.

## Results

### Identification of genes derived from molecular domestication of TEs

In order to detect new transposon domestication events in the mammalian lineage, we looked for protein-coding genes made up by TE-derived sequences. By identifying genes with a coding sequence overlapping greater than 50% with annotated TEs and present in more than one species, we obtained a list of 28 and 9 candidates in the human and mouse genomes, respectively (Additional file 1: Table S1). We recovered well-known TE-derived genes, such as *syncytin-1* [14], *syncytin-2* [15], *SETMAR* [16], and the *ZBED* [17] and *PNMA* families [18]. More ancient domesticated transposons like the vertebrate-specific *RAG1* and *RAG2* genes [3] were not detected, probably due to higher sequence divergence in relation to the ancestral TE element. We also found new

putatively domesticated TEs, most with unknown function, confined to either primate or *Mus* species. However, our attention was drawn to the remarkable case of *Tceal7*, a gene present in all major groups of placental mammals. Due to its broad phylogenetic range, we decided to study this molecular domestication event in further detail.

### Molecular composition of the *Tceal7* gene

*Tceal7* is a small gene consisting of two 5′ non-coding exons and a third exon that includes the whole open reading frame (ORF) and the 3′UTR. We observed an overlap of 76% between the *Tceal7* ORF sequence and that of HAL1b, a non-LTR retrotransposon belonging to the long interspersed nuclear element-1 (LINE-1 or L1) superfamily (Fig. 1a). Moreover, the last 18 nucleotides of the *Tceal7* ORF and most of the 3′UTR originated from two other L1 subfamilies: the elements L1MEe and L1ME4a (Fig. 1a and Additional file 1: Fig. S1). We observed a partial retention of the original TE coding frame, as human TCEAL7 amino acid sequence shares 31.7% identity with the ORF1p of HAL1b (Fig. 1b). From these results, we determined that the *Tceal7* gene arose from a composite sequence derived from two L1 elements (Fig. 1).

Although *Tceal7* stands out as the most similar to the ancestral L1 retrotransposon sequences, it is but a representative member of a multigenic family called *Bex/Tceal*. This family forms a gene cluster on the X chromosome of placental mammals, with no detectable orthologs outside this clade [21]. In humans, the cluster consists of 5 *Bex* (brain-expressed X-linked) and 9 *Tceal* (transcription elongation factor A (SII)-like) genes, spanning ~1.5 Mb. In mouse, the number is reduced to 11 genes due to the lineage-specific loss of *Tceal2*, *Tceal4*, and *Bex5*. Interestingly, we detected the presence of clustered *Bex/Tceal* genes in each major eutherian lineage (Additional file 1: Fig. S2), confirming that the origin and expansion of the family took place after the divergence of the marsupial-placental clades, and before the radiation of the latter. In agreement, HAL1b, L1MEe, and L1ME4a elements have been suggested to be active retrotransposons during the appearance of early eutherians ~150 Mya [22].

### Evolutionary diversification of the *Bex/Tceal* family

BEX and TCEAL protein sequences are relatively divergent [21], likely due to low selective constraints after the retrotransposon domestication. However, both families share a conserved region toward their C-terminal end (Additional file 1: Fig. S3A). BEX proteins, with the exception of BEX4, are predicted to have a coiled coil within this region [23–25], and we found that TCEAL1, TCEAL7, and TCEAL8 are also predicted to harbor a C-terminal coiled coil domain (Additional file 1: Fig. S3A). Remarkably, most of the protein-protein interactions of BEX1, BEX2, and BEX3 have been mapped to this domain [24], highlighting its potential functional relevance. Furthermore, BEX proteins have been classified as intrinsically disordered proteins (IDPs) [24]. We found that TCEAL proteins, as previously described for BEX members [24], are predicted to have a disordered N-terminal region and C-terminal α-helices (Additional file 1: Fig. S4). Interestingly, we also detected all these structural features in the ancestral transposon HAL1b (Additional file 1: Fig. S4), suggesting that these features were preserved along the domestication process and inherited by the *Bex/Tceal* genes.
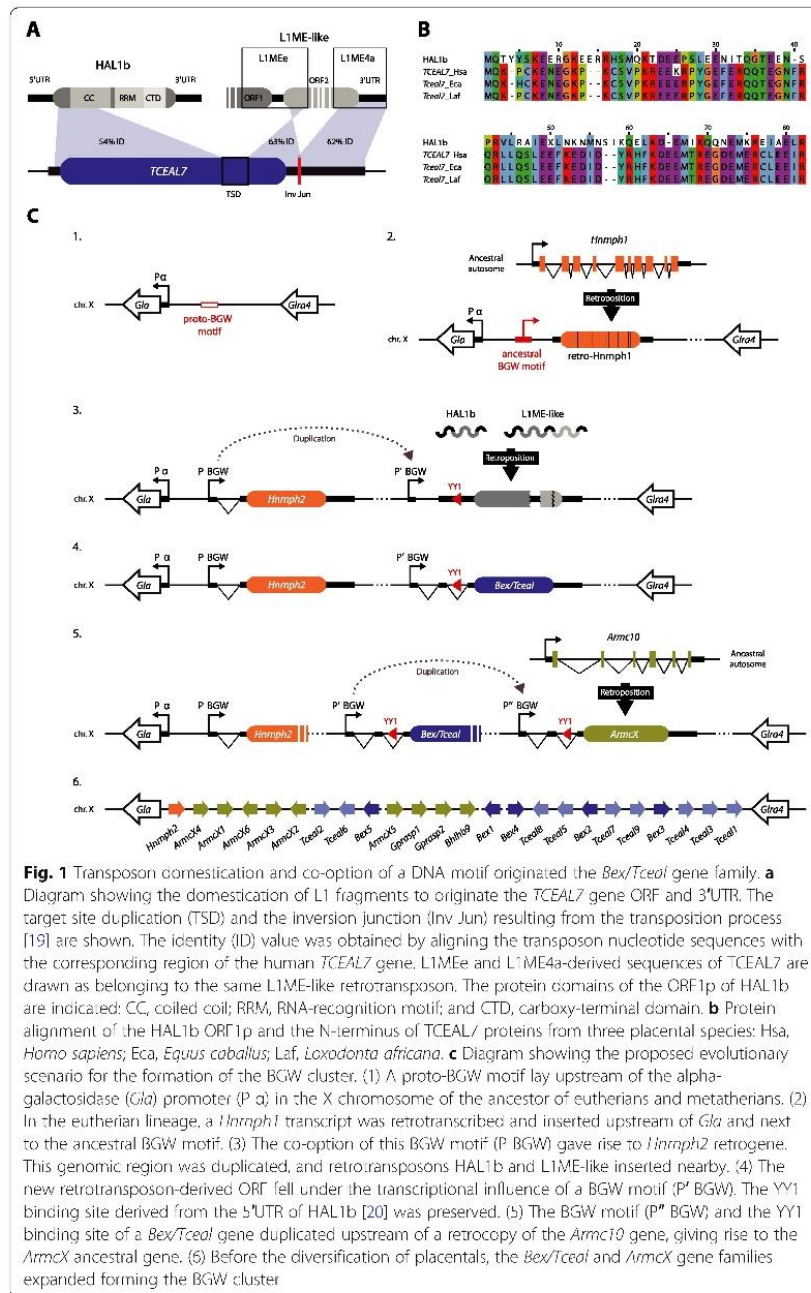
**Fig. 1** Transposon domestication and co-option of a DNA motif originated the *Bex/Tceal* gene family. **a** Diagram showing the domestication of L1 fragments to originate the *TCEAL7* gene ORF and 3'UTR. The target site duplication (TSD) and the inversion junction (Inv Jun) resulting from the transposition process [19] are shown. The identity (ID) value was obtained by aligning the transposon nucleotide sequences with the corresponding region of the human *TCEAL7* gene. L1MEe and L1ME4a-derived sequences of TCEAL7 are drawn as belonging to the same L1ME-like retrotransposon. The protein domains of the ORF1p of HAL1b are indicated: CC, coiled coil; RRM, RNA-recognition motif; and CTD, carboxy-terminal domain. **b** Protein alignment of the HAL1b ORF1p and the N-terminus of TCEAL7 proteins from three placental species: Hsa, *Homo sapiens*; Eca, *Equus caballus*; Laf, *Loxodonta africana*. **c** Diagram showing the proposed evolutionary scenario for the formation of the BGW cluster. (1) A proto-BGW motif lay upstream of the alpha-galactosidase (*Gla*) promoter (P α) in the X chromosome of the ancestor of eutherians and metatherians. (2) In the eutherian lineage, a *Hnrnph1* transcript was retrotranscribed and inserted upstream of *Gla* and next to the ancestral BGW motif. (3) The co-option of this BGW motif (P BGW) gave rise to *Hnrnph2* retrogene. This genomic region was duplicated, and retrotransposons HAL1b and L1ME-like inserted nearby. (4) The new retrotransposon-derived ORF fell under the transcriptional influence of a BGW motif (P' BGW). The YY1 binding site derived from the 5'UTR of HAL1b [20] was preserved. (5) The BGW motif (P" BGW) and the YY1 binding site of a *Bex/Tceal* gene duplicated upstream of a retrocopy of the *Armc10* gene, giving rise to the *ArmcX* ancestral gene. (6) Before the diversification of placentals, the *Bex/Tceal* and *ArmcX* gene families expanded forming the BGW cluster

Previous studies using a reduced number of mammalian species reported gene conversion events and positive selection signatures on some *Bex/Tceal* genes [21, 26]. We decided to perform an expanded analysis adding species from the major eutherian clades, and found homogenization of coding sequences among three

Navas-Pérez *et al. Genome Biology*      (2020) 21:267

Page 5 of 27

groups of *Bex/Tceal* paralogs (*Tceal2* with *Tceal4*; *Tceal3* with *Tceal5* and *Tceal6*; and *Bex1* with *Bex2*) across all studied lineages (Additional file 1: Fig. S3B). Moreover, after filtering out sequences experiencing gene conversion [27], we detected several sites with signatures of positive selection, mainly in the *Bex* subfamily tree, and an episode of positive selection after the branching of *Bex5* (Additional file 1: Fig. S5). This suggests that an ancestral gene within the branch leading to *Bex3* and *Bex4* went through an adaptive process in an eutherian ancestor before the diversification of placental mammals.

### Neighboring DNA sequence co-option as a central regulatory element

Within the same eutherian-specific chromosomal region containing *Bex* and *Tceal* genes, there are multiple retrocopy genes belonging to another gene family (the *Armcx* family, also known as *ALEX*) [28]. Despite no protein similarity between both families, they share a homologous DNA sequence motif in their promoter region known as the BGW motif [21], a ~ 60 base pair (bp)-long sequence containing an internal E-box, which has been shown to be essential for the regulation of mouse *Tceal7* [29] and human *ARMCX1* [30] expression. The promoter region of *Hnrnph2*, another eutherian-specific retrogene located at the centromeric end of the cluster, also harbors this unique motif [21]. This multiplicity of BGW motifs raises the question of how genes with three independent origins ended up with separate, but homologous, regulatory elements. The promoter of *Hnrnph2* is bidirectional and shared with the galactosidase-alpha (*GLA*) gene [31]. We found a BGW-like motif upstream of the *GLA* promoter in marsupials that lacks an eutherian-restricted 11 bp sequence required for the proper transcription of human *ARMCX1* gene [30] (Additional file 1: Fig. S6). Therefore, the origin of the BGW motif can be traced back to sequences already present in the *GLA* promoter of the last therian common ancestor. Although we cannot determine the precise order of the events leading to the assembly of the whole cluster, the inferred co-option of the BGW motif by the ancestors of the *Bex/Tceal* and *Armcx* families allows us to reconstruct the main steps of this evolutionary process (diagrams depicted in Fig. 1c, see legend for details).

### Brain expression of the *Bex/Tceal* genes and deregulation in neuropsychiatric disorders

Information about expression patterns for *Bex/Tceal* genes is disperse, with data restricted to human, mouse, or rat [12, 29, 32–37]. We gathered publicly available transcriptomic data from eight homologous adult organs of five species belonging to main placental lineages. We observed that most genes present a tissue-enriched pattern, with brain being the organ showing the highest expression levels for most paralogs across species (Fig. 2a). Although brain is an organ where many genes tend to be expressed [38], this result suggests that some of the neural functions reported in mouse and human for this gene family [11, 12, 39] might be conserved among the eutherian clade.

We also investigated the expression patterns of this group of genes during mouse development by in situ hybridization and observed a subset of *Bex/Tceal* genes being highly and widely expressed during mouse embryogenesis, especially *Bex* genes (Fig. 2b and Additional file 1: Fig. S7). *Bex3* expression was particularly strong in

the central nervous system during murine development compared to other members of the family (Fig. 2b).

By analyzing publicly available human transcriptomic data of autism spectrum disorder (ASD) and schizophrenia, two well-studied neuropsychiatric disorders, we found a significant decrease in the expression of *BEX/TCEAL* genes in patients compared to controls in different brain regions and datasets (Additional file 1: Table S2), being *BEX* but not *TCEAL* genes significantly enriched among the differentially expressed genes in most datasets (Additional file 1: Table S3). Notably, *BEX3* is located in the interval associated with the neurological features of patients diagnosed with early-onset neurological disease trait (EONDT), which harbor different genomic deletions encompassing *BEX/TCEAL* genes [40–42].

### Activation of neural tube progenitor proliferation by BEX/TCEAL proteins in a non-eutherian vertebrate

Next, we investigated the putative function of the inferred, original TE composite sequence that later evolved into the *Bex/Tceal* family. To understand the potential physiological response of this ancient element in the eutherian ancestor, we aimed to mimic the original scenario by using a non-eutherian organism. For this purpose, we synthetically reconstructed a version of the ancestral *Bex/Tceal* protogene, termed *HALEX* (**HAL**1b-derived on **e**utherian **X** chromosome), based on TE consensus sequences (see "Methods"). Moreover, we also studied two *Bex/Tceal* members: *Tceal7*, the gene with the highest sequence similarity to the original protogene; and *Bex3*,
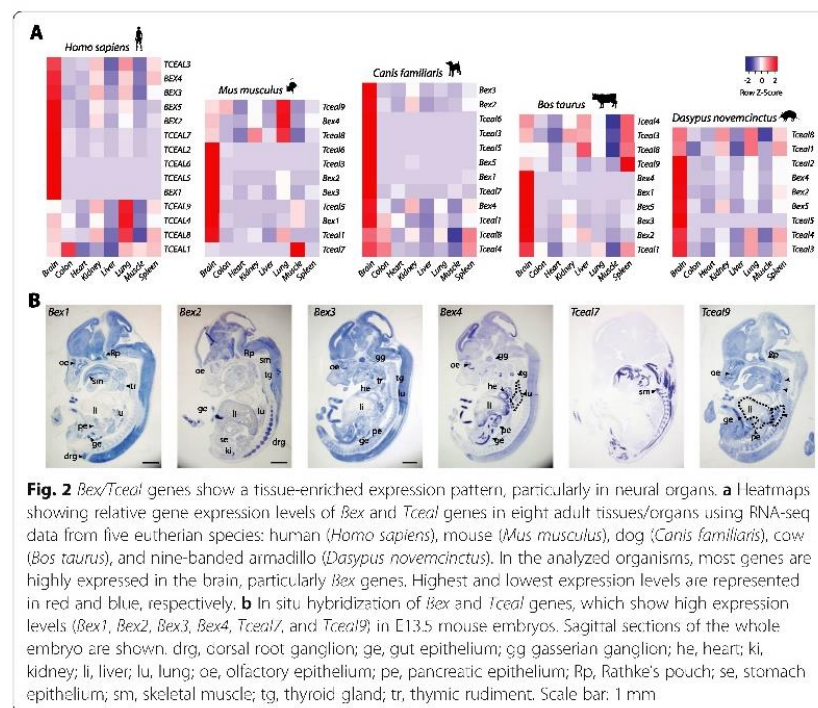


**Fig. 2** *Bex/Tceal* genes show a tissue-enriched expression pattern, particularly in neural organs. **a** Heatmaps showing relative gene expression levels of *Bex* and *Tceal* genes in eight adult tissues/organs using RNA-seq data from five eutherian species: human (*Homo sapiens*), mouse (*Mus musculus*), dog (*Canis familiaris*), cow (*Bos taurus*), and nine-banded armadillo (*Dasypus novemcinctus*). In the analyzed organisms, most genes are highly expressed in the brain, particularly *Bex* genes. Highest and lowest expression levels are represented in red and blue, respectively. **b** In situ hybridization of *Bex* and *Tceal* genes, which show high expression levels (*Bex1*, *Bex2*, *Bex3*, *Bex4*, *Tceal7*, and *Tceal9*) in E13.5 mouse embryos. Sagittal sections of the whole embryo are shown. drg, dorsal root ganglion; ge, gut epithelium; gg gasserian ganglion; he, heart; ki, kidney; li, liver; lu, lung; oe, olfactory epithelium; pe, pancreatic epithelium; Rp, Rathke's pouch; se, stomach epithelium; sm, skeletal muscle; tg, thyroid gland; tr, thymic rudiment. Scale bar: 1 mm

which shows the strongest expression in the embryonic nervous system. We electroporated murine *Bex3*, *Tceal7*, and *HALEX* genes into the neural tube of chicken embryos at stage HH12 to investigate their capacity to elicit a cellular response in neuronal progenitors. Expression of *Bex3* and *Tceal7*, but not *HALEX*, generated a significant increase in cell proliferation in the chicken embryonic neural tube, similarly to previous reports in mammalian cultured cells [11, 43] (Additional file 1: Fig. S8). Although the heterologous approach we have used cannot reproduce the regulatory and signaling environments of the eutherian ancestor, the observation that only the murine constructs, but not the ancient element, produced a measurable response suggests that the current ability of *Bex/Tceal* genes to effectively modulate cellular physiology was not present in the ancestral protogene, being acquired during eutherian evolution.

### Generation of two independent *Bex3* mutant lines

While functional studies for the *Bex/Tceal* genes have focused mainly on in vitro assays, little is known about their role at the organism level. Based on previous reports linking *Bex3* to neuronal physiology [11, 43, 44], the analyses on patients with neurological features harboring deletions encompassing *BEX/TCEAL* genes [40–42], and its neural-enriched expression in adult and embryonic tissues, we decided to generate mouse mutant lines for this gene. We used CRISPR-Cas9 technology to generate mutant alleles for *Bex3* in mice and selected two for in-depth characterization (Fig. 3a and Additional file 1: Fig. S9). One of them, namely $Bex3^{KO}$, carried a 196-bp deletion that caused a frameshift mutation, introducing a premature stop codon that led to a truncated coding sequence. The second line, which we named $Bex3^{\Delta(24-72)}$, carried a 147-bp deletion that removed 49 amino acids of the central core of the protein, specifically the pro-apoptotic domain, and retained the C-terminal coiled coil domain required for BEX3 dimerization, nuclear import, ubiquitination [44], and interaction with its multiple partners [24]. Therefore, this mutation potentially expresses a protein lacking an essential functional domain, which is likely to act as a hypomorph or a dominant negative allele, as it has been shown to be the case in cell culture experiments [44].

### Anatomical alterations in skull and brain of *Bex3* mutant mice

Homozygous mutant mice for both lines could be distinguished from wild-type counterparts by external observation of subtle facial differences (Fig. 3b, c), hence skull measurements were taken out in order to identify the origin of these alterations. Snout-to-midbone length (D2 in Additional file 1: Fig. S10) tended to be smaller in the mutants, while the width of the eye sockets (D24 in Fig. 3c) was significantly larger in $Bex3^{KO}$ mice. Further, the ratio between these measurements was reduced by 19.4% ± 3.4 in $Bex3^{KO}$ ($p = 0.001$) and 10.2% ± 3.4 in $Bex3^{\Delta(24-72)}$ ($p = 0.064$), suggesting abnormal frontal bone morphology that resulted in a rounder eye appearance. Additionally, lack of fully functional *Bex3* also led to increased size of posterior parameter measurements (D5, D19-D21 in Fig. 3c and Additional file 1: Fig. S10), indicating that the bone cavity harboring the cerebellum was larger in the mutants, particularly in $Bex3^{KO}$ mice.

The overall brain and cerebellum anatomical structures of $Bex3^{KO}$ and $Bex3^{\Delta(24-72)}$ mice appeared close to normal, although they showed a reduction in cortical surface and size, as well as in cerebellum size (Fig. 3d, e) despite the layered

**Fig. 3** CRISPR-Cas9-generated *Bex3* mutant alleles show skull and brain abnormalities. **a** Schematic representation of the mouse *Bex3* locus (mm10; chrX:136,270,126-136,272,051) depicting exon-intron structure (non-coding and coding exons represented as black and blue rectangles, respectively). Two sgRNAs (red arrows) were used to generate CRISPR-Cas9-mediated deletions. The altered proteins potentially produced by the edited alleles that were selected to generate homozygous mice are shown below. In the *Bex3^{KO}* line, the deletion caused a frameshift in the open reading frame (region in gray) leading to the appearance of a premature STOP codon. The deletion in the *Bex3^{Δ(24−72)}* line removed 54 amino acids from the central core of the protein. AD, pro-apoptotic domain; CC, coiled coil domain; NES, nuclear export signal. **b, c** Morphometric analyses of skulls from wild-type and mutant *Bex3^{KO}* 6-week-old males employing a total of 24 anatomical measurements **b** revealed that *Bex3* dysfunction led to cranial abnormalities in frontal bone and skull height **c** (the complete analysis can be found in Additional file 1: Fig. S10). Measurements were normalized to maximum skull length (D0) and expressed relative to controls (black horizontal line). Deviations of 5% with respect to controls are shown as dotted red and green horizontal lines. Results are presented as mean ± SEM ($N ≥ 4$); *$p < 0.05$, one-way ANOVA. **d, e** *Bex3* mutant brains showed altered brain morphology as evidenced by gross **d** and refined **e** anatomical measurements, concomitant with enlarged ventricular surfaces. Scale bar: 1 mm. Results in **d** and **e** are presented as mean ± SEM ($N ≥ 5$); *$p < 0.05$, **$p < 0.01$, ***$p < 0.005$, ****$p < 0.001$, one-way ANOVA

structures of the cortex, hippocampus, and cerebellum, being preserved in the mutants (data not shown). In addition, we observed a marked increase in brain ventricular surface in both mutant lines (Fig. 3e). It is worth noting that reduced brain size and enlargement of brain ventricles have been described extensively in neurodevelopmental diseases [45–47].

### Behavioral defects in *Bex3* mutant mice

To determine if the anatomical and physiological brain defects observed in the *Bex3* mutant mice have an effect on animal behavior, adult mutant and control mice were subjected to a comprehensive battery of behavioral tests (Fig. 4 and Additional file 1: Fig. S11). In non-stressful open field, *Bex3* mutant mice showed normal locomotor activity [F $(10,1) = 0.291$; $p = 0.601$ and F $(10,1) = 1.628$; $p = 0.230$] for *Bex3^{KO}* and *Bex3^{Δ(24−72)}* lines respectively (Fig. 4a). However, we observed that both *Bex3^{KO}* and *Bex3^{Δ(24−72)}* lines displayed significantly more repetitive behavior events [rearing: *Bex3^{KO}*: F $(10,1) = 5.078$; $p =$

0.047 and $Bex3^{\Delta(24-72)}$: F (10,1) = 1.355; $p$ = 0.271; grooming: $Bex3^{KO}$: F (10,1) = 6.188; $p$ = 0.032 and $Bex3^{\Delta(24-72)}$: F (10,1) = 22.216; $p < 0.001$] when compared to their wild-type littermates (Fig. 4a). We evaluated social interaction using a three-chambered assay to study the interactions with familiar or stranger mice (Fig. 4b). While wild-type mice showed preferential interaction with familiar mice against the empty compartment [F (10, 1) = 94.218; $p < 0.001$], we only found impairment in the interaction with the familiar mice in $Bex3^{KO}$ line, with no significant differences in the sniffing time between empty and the familiar mice containing compartments [F (10,1) = 2.053; $p$ = 0.182]. In contrast, in a social novelty paradigm, control animals spent significantly more time in close interaction with the novel animal, whereas $Bex3^{KO}$ displayed no significant preference for social novelty [F (10,1) = 3.114; $p$ = 0.108] or even avoided social novel interaction in $Bex3^{\Delta(24-72)}$ [F (10, 1) = 51.494; $p < 0.001$]. Furthermore, only $Bex3^{KO}$ mice also showed impairment in nest building [F (10,1) = 9.800; $p$ = 0.010; Fig. 4c], which has been correlated with abnormal social organized behavior [48, 49]. Sensorimotor gating, whose impairment is also strongly associated with some neurodevelopmental disorders [50], was assayed by prepulse inhibition (PPI) of acoustic startle reflex. Although $Bex3$ mutants showed similar acoustic startle reflex than wild-type mice (Additional file 1: Fig. S11), both mutant lines exhibited a significant enhancement of PPI [$Bex3^{KO}$: F (10,1) = 6.588; $p$ = 0.028 and $Bex3^{\Delta(24-72)}$: F (10,1) = 6.038; $p$ = 0.033; Fig. 4d]. Moreover, cognitive behavior was assayed in the Y-maze test, object recognition memory, and passive avoidance tests. In the Y-maze test, $Bex3$ mutant mice showed significantly lower spontaneous alternation indexes, indicative of attention deficits and/or working memory [$Bex3^{KO}$: F (10,1) = 5.186; $p$ = 0.045 and $Bex3^{\Delta(24-72)}$: F (10,1) = 8.125; $p$ = 0.017; Fig. 4e]. On the contrary, only $Bex3^{KO}$ mice showed significant lower performance compared to their control littermates in object recognition memory [$Bex3^{KO}$: F (20,3) = 4.843; $p$ = 0.039 and $Bex3^{\Delta(24-72)}$: F (20,3) = 0.549; $p$ = 0.467, analyzed by 2-way ANOVA for genotype X session interaction; Fig. 4f] and passive avoidance tests [$Bex3^{KO}$: F (20,3) = 6.548; $p$ = 0.018 & $Bex3^{\Delta(24-72)}$: F (20,3) = 0.262; $p$ = 0.613, analyzed by 2-way ANOVA for genotype X session interaction; Fig. 4g]. In summary, the behavioral experiments indicated that both $Bex3$ mutant lines display repetitive behaviors and abnormal social conducts and that $Bex3^{KO}$ mice presented more severe phenotypes, especially in cognitive tests.

### Reduction in the number of cortical and subcortical interneurons in *Bex3*-deficient mice

A link between the disruption of interneuron inhibitory circuits in the neocortex and some clinical aspects of neurodevelopment and psychiatric disorders has already been established [51]. Our data show that both $Bex3$ mutant lines display gross brain morphological alterations and have abnormal behavioral traits and that $Bex3^{KO}$ mice presented more severe phenotypes, especially in terms of memory and learning impairment. These results, together with transcriptomic and structural genomic data related to human neurological disorders (Additional file 1: Tables S2 and S3) [40–42], suggest that $Bex3$ may function through the maintenance/renewal of specific neurons, and its absence can give rise to neurological conditions.

We measured parvalbumin-positive (PV+) interneuron density and found it was significantly decreased in motor, cingulate and sensory cortices, and caudate-putamen of $Bex3^{KO}$ mice; $Bex3^{\Delta(24-72)}$ mice showed a significant reduction of PV+ interneurons

**Fig. 4** *Bex3* mutant mice display an array of behavioral and cognitive alterations. Adult control and *Bex3* mutant mice were subjected to a comprehensive battery of behavioral tests. **a** Locomotor activity (number of broken beams) and total repetitive behaviors (number of grooming and rearing) were evaluated in a 5-min open field test. **b** Social behavior was assayed in the three-chambered test. Sniffing times are showed in social familiar and social stranger phases. **c** Social organized behaviors were tested by the nesting building assay. **d** Sensory motor integration was determined by prepulse inhibition of acoustic reflex. **e** The Y-maze test was performed to evaluate repetitive conducts, attention, and/or working memory. As alternance, the % of triplets was determined. **f** Novel object recognition memory was evaluated 24 h after a training session. A positive discrimination index in this memory test indicates novel object preference exploration. **g** The passive avoidance test was used to evaluate emotional cognition. Entry latencies to the dark compartment in training and memory testing sessions are shown. Results are presented as mean ± SEM ($N = 6$ mice per experimental group); *$p < 0.05$, **$p < 0.01$, ***$p < 0.005$, ****$p < 0.001$, one-way ANOVA

only in cingulate and sensory cortices (Fig. 5a, b) and minor, non-significant, decrease in motor cortex and caudate-putamen. On the other hand, altered numbers of hippocampal neurons and interneurons have been associated to the onset of social and cognitive deficits in several animal models for neurological disorders (reviewed in [52]). Even though the hippocampus-to-hemisphere size ratio was not significantly altered in *Bex3* mutant mice (Fig. 5c), we observed a strong reduction in the density of total neurons and, particularly, calretinin-positive (CR+) inhibitory interneurons in the *stratum radiatum* of the CA1–2 hippocampal fields of $Bex3^{KO}$ mice (Fig. 5d, e).

We decided to analyze the status of the adult hippocampal neurogenic niche in *Bex3* mutant mice. Interestingly, the density of immature neurons was significantly decreased in the subgranular zone of $Bex3^{KO}$ mice, but not in $Bex3^{\Delta(24-72)}$ (Fig. 5f, g), thus suggesting that adult neurogenesis could be partially compromised in these animals.

**Alteration in the excitation/inhibition balance in the hippocampal CA2 region of $Bex3^{KO}$ mice**
Impaired social interaction and social learning is known to rely on the CA2 neuronal circuit within the hippocampus [53, 54], while altered interneuron numbers in the CA2 hippocampal field can disrupt the excitatory/inhibitory hippocampal balance. Thus, we

tested if CA2 synaptic transmission was altered in *Bex3^{KO}* mice, which have reduced numbers of hippocampal interneurons in this area. To determine whether the evoked synaptic activity was altered in hippocampal pyramidal CA2 neurons, we monitored evoked excitatory (eEPSCs) and inhibitory (eIPSCs) postsynaptic currents. We found that the amplitude of eIPSCs—but not eEPSCs—was clearly decreased in mutant mice (Fig. 5h). However, no statistically significant differences were found between wild-type and *Bex3^{KO}* mice in the input-output curve, indicating that the net basal synaptic activity was not affected in these animals (Fig. 5i). Additionally, we studied spontaneous synaptic transmission onto hippocampal CA2 pyramidal neurons (Fig. 5j-m). Total spontaneous activity frequency (sEPSC + sIPSC) was significantly decreased in *Bex3^{KO}* mice respect to wild-type (Fig. 5l) due to a strong reduction in sIPSCs frequency. On the other hand, no differences were found between sEPSCs and sIPSCs amplitude when comparing wild-type and mutant mice (Fig. 5m). These results indicate an excitation/inhibition imbalance in the hippocampal CA2 region of *Bex3^{KO}* mice due to a strong decrease in inhibitory synaptic transmission, which is in agreement with the reduced density of inhibitory interneurons observed in this region.

### Altered mTOR signaling in the adult brain of *Bex3*-deficient mice

BEX3 protein physically interacts with the TSC1/TSC2 complex [55], which is essential for the proper regulation of the mTOR signaling cascade [56]. This pathway controls brain development and function at multiple levels, and its dysregulation has been implicated in several neurological diseases [57]. Intriguingly, we found a decrease in the phosphorylation of the mTORC1 readout S6K1 in brain lysates of both *Bex3* mutant lines, which would suggest mTORC1 hypoactivation; in contrast, no changes were detected in other targets like 4E-BP1 or S6. The mTORC2 readouts AKT and NDRG1 were hyperphosphorylated, whereas PKCα and PKCγ protein levels were unaltered (Fig. 6 and Additional file 1: Fig. S12). All these data suggest a hyperactivation of mTORC2 signaling in the adult brain as one of the possible molecular consequences linked to *Bex3* deficiency.

### Discussion

By screening the human and mouse genomes for domesticated transposons, we have identified the footprints of two L1 retrotransposons within the human *TCEAL7* gene, indicating that a domestication event restricted to placental mammals gave rise to a whole new gene family in the X chromosome. After the eutherian-specific gene L1TD1 [58], the *Bex/Tceal* family represents the second case of domestication of non-LTR transposons reported in metazoans, and the first to give rise to a multigenic family. This domestication generated an ancestral protogene that we termed *HALEX*, and our experiments in a non-eutherian model suggest an evolutionary scenario where this element was progressively integrated into anciently established gene networks through complete neofunctionalization before the diversification of eutherians. This is in contrast with most other known domestication cases, where new genes perform tasks at the cellular level similar to those of the ancestral element [59]. Several members of the *Bex/Tceal* gene family have been reported to code for hub proteins, due to their high number of interactions with multiple proteins belonging to several signaling pathways
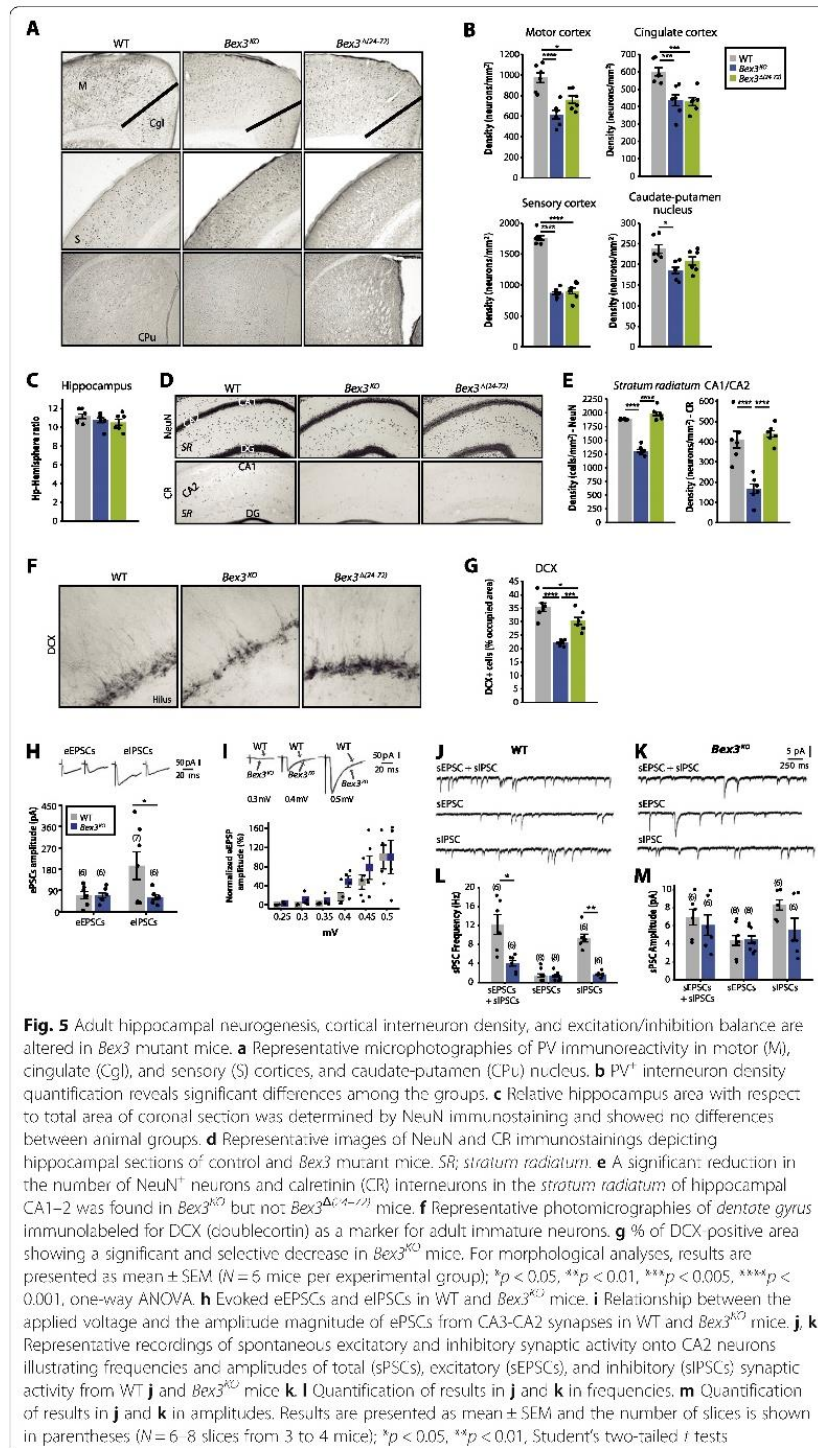
**Fig. 5** Adult hippocampal neurogenesis, cortical interneuron density, and excitation/inhibition balance are altered in *Bex3* mutant mice. **a** Representative microphotographies of PV immunoreactivity in motor (M), cingulate (Cgl), and sensory (S) cortices, and caudate-putamen (CPu) nucleus. **b** PV⁺ interneuron density quantification reveals significant differences among the groups. **c** Relative hippocampus area with respect to total area of coronal section was determined by NeuN immunostaining and showed no differences between animal groups. **d** Representative images of NeuN and CR immunostainings depicting hippocampal sections of control and *Bex3* mutant mice. *SR; stratum radiatum*. **e** A significant reduction in the number of NeuN⁺ neurons and calretinin (CR) interneurons in the *stratum radiatum* of hippocampal CA1–2 was found in *Bex3^{KO}* but not *Bex3^{Δ(24–71)}* mice. **f** Representative photomicrographies of *dentate gyrus* immunolabeled for DCX (doublecortin) as a marker for adult immature neurons. **g** % of DCX positive area showing a significant and selective decrease in *Bex3^{KO}* mice. For morphological analyses, results are presented as mean ± SEM ($N = 5$ mice per experimental group); $*p < 0.05$, $**p < 0.01$, $***p < 0.005$, $****p < 0.001$, one-way ANOVA. **h** Evoked eEPSCs and eIPSCs in WT and *Bex3^{KO}* mice. **i** Relationship between the applied voltage and the amplitude magnitude of ePSCs from CA3-CA2 synapses in WT and *Bex3^{KO}* mice. **j**, **k** Representative recordings of spontaneous excitatory and inhibitory synaptic activity onto CA2 neurons illustrating frequencies and amplitudes of total (sPSCs), excitatory (sEPSCs), and inhibitory (sIPSCs) synaptic activity from WT **j** and *Bex3^{KO}* mice **k**. **l** Quantification of results in **j** and **k** in frequencies. **m** Quantification of results in **j** and **k** in amplitudes. Results are presented as mean ± SEM and the number of slices is shown in parentheses ($N = 6$–8 slices from 3 to 4 mice); $*p < 0.05$, $**p < 0.01$, Student's two-tailed *t* tests

[11, 12, 35, 60–64]. Remarkably, we found that both BEX/TCEAL proteins and the ancestral HAL1b sequence from which they derive are predicted to be predominantly disordered at their N-terminus and contain C-terminal α-helices (Additional file 1: Fig. S4). Relatedly, disorder is a structural property that frequently translates into variable conformations and the ability to bind multiple partners [65]. Thus, we suggest that the evolutionary potential of *Bex/Tceal* genes for a multiplicity of protein-protein interactions might be a direct legacy of the ancestral HAL1b retrotransposon protein.

The mammalian target of rapamycin (mTOR) pathway stands out as one of the molecular networks in which *Bex/Tceal* genes integrated, as shown by previous work on *BEX2* and *BEX4* [62, 64]. The mTOR protein kinase is involved in a myriad of processes related to cell growth, proliferation, and survival. It can assemble into two molecularly and functionally different complexes that include some specific components such as mTORC1 and mTORC2 [66]. mTORC1 main substrates are S6K and 4E-BP kinases, while mTORC2 downstream targets include AKT and SGK kinases. Upon phosphorylation, AKT inhibits the TSC1/2 complex [67], which regulates both mTORC signaling cascades in opposite directions: it inhibits mTORC1 and activates mTORC2. Yet, it can interact physically with mTORC2 only, specifically through Tsc2 [67].

Our data show that *Bex3* deficiency leads to hyperphosphorylation of Akt on Ser-473 in brain extracts, suggesting mTORC2 hyperactivation. Akt hyperphosphorylation should also induce hyperactivation of the mTORC1 route but no changes in 4E-BP1 phosphorylation were detected, while S6K1 was found to be hypophosphorylated on Thr-389. Importantly, earlier phosphorylation steps on the autoinhibitory domain of S6K1 are needed to render residue Thr-389 accessible to mTORC1. Both mTORC1 and JNK are involved in the phosphorylation of this domain of S6K1 in muscle [68]. JNK is inhibited by mTORC2, and therefore, mTORC2 hyperactivation in our models could indirectly lead to the observed S6K1 hypophosphorylation on Thr-389. Surprisingly, phosphorylation levels of S6K1 main effector S6 on residues Ser235/236 were not affected, suggesting the possible existence of compensatory mechanisms by other kinases that can target these residues [69]. In summary, our biochemical data and the fact that BEX3 interacts with the TSC1/TSC2 complex [55] suggest the putative involvement of the mTOR pathway in some of the reported phenotypic alterations. However, the weak differences found in some mTOR effectors imply that additional signaling pathways may also be involved in the reported changes.



**Fig. 6** *Bex3* deficiency leads to aberrant mTOR signaling in the brain. **a** Western blot analyses of whole brain lysates of adult *Bex3^KO* mice revealed abnormal phosphorylation ratios of some mTORC1 and mTORC2 targets. Representative images. **b** Quantification of data in **a**, relative to wild-type. Results are presented as mean ± SEM ($N = 4$–5 per experimental group); $*p < 0.05$, non-parametric Mann–Whitney test

Given that BEX3 physically interacts with TSC1, an interaction shown to prevent proteasome degradation and thus necessary for NGF-p75NTR-BEX3-mediated apoptosis [55], we propose that BEX3 could prevent the TSC1/2 complex from interacting with mTORC2, eventually reducing the activity of the pathway. The absence of *Bex3* would then cause a hyperactivation of the mTORC2 pathway without affecting mTORC1. Thus, the function of BEX3 could be that of fine-tuning the regulation of these cascades. The fact that the $Bex3^{KO}$ and $Bex3^{\Delta(24-72)}$ lines show analogous results suggests that the missing central core of the mutant $Bex3^{\Delta(24-72)}$ protein is necessary for TSC1/2 recognition or, alternatively, that its coupling with the TSC1/2 complex does not interfere with mTORC2 binding, presumably due to the smaller size of the resultant protein. Furthermore, since mTOR signaling has been involved in adult neurogenesis and neuronal excitability and is dysregulated in several neurological diseases, the mTOR pathway is a compelling candidate to explain, at least in part, the molecular alterations behind the phenotype observed in *Bex3*-deficient mice.

The acquisition of regulatory sequences enabling transcription is also a crucial but scarcely studied step during the process of domestication and subsequent neofunctionalization of transposable elements [70]. We have described here a short, non-coding regulatory region called BGW motif that was co-opted during eutherian evolution by three unrelated gene ancestors generated from retrotransposition events.

Although transposons and mammalian-restricted genes predominantly show tissue-specific patterns [71, 72], we observed a relatively wide expression for most *Bex/Tceal* genes. In mouse embryos, expression profiles of *Bex/Tceal* genes are consistently associated with proliferative tissues within organs such as the stomach, lungs, pancreas, or central nervous system. Interestingly, some of these genes are used as markers for progenitor cells in developing tissues and are involved in cell proliferation, differentiation, and cell death [11, 29, 34, 73, 74]. In adult tissues, our results suggest that *Bex3* regulates adult hippocampal neurogenesis, while some *Bex/Tceal* genes are induced upon injury, having an impact on axonal, muscular, and hepatic regeneration [29, 37, 39].

*BEX1* and *BEX3* have been recently identified among the most significantly downregulated genes in excitatory neurons of the prefrontal cortex of Alzheimer's disease patients [75]. Furthermore, patients with severe intellectual disability, craniofacial dysmorphism, and autism have been reported to carry different genomic microdeletions in Xq22 encompassing *BEX/TCEAL* genes, with *BEX3* pinpointed as one of the main candidates to cause these neurological features [40–42]. Also, a small 252-kb duplication spanning *BEX3, TCEAL4, TCEAL9,* and *RAB40A* has been reported in a patient with autism (Decipher database, ID: 290829) [76]. Here we show that mutations of murine *Bex3* lead to subtle craniofacial changes and have a profound impact on repetitive and social behavioral performance, two important behavior alterations required to diagnose autism spectrum disorders (ASD) [77]. Consistent with ASD-like behaviors, some cerebral cortex areas, related with social behaviors [78, 79], and the striatum, related with repetitive behavior [80], showed alterations in the number of parvalbumin interneurons, another feature found in ASD mice models and patients [81–88]. Cognitive alterations are frequently found in ASD patients [89]. In this aspect, only $Bex3^{KO}$ mice show learning and memory defects in object recognition and passive avoidance tests, two hippocampus-dependent paradigms [90–92]. In parallel, alterations in calretinin interneurons in the CA1–2 hippocampal *stratum radiatum*, excitatory-inhibitory

imbalance in the CA2 field, and potential defects in adult *dentate gyrus* neurogenesis were also found only in $Bex3^{KO}$ mice. In this regard, the less severe phenotype of $Bex3^{\Delta(24-72)}$ mice suggests that the resulting protein may act as a hypomorph version of the wild-type allele. In brief, the phenotypical features of *Bex3* mutant mice makes this gene an exciting candidate for future research into human neurological disorders that impact upon repetitive behavior, sociability, and intellectual disability.

## Conclusions

We characterized the evolutionary process by which an ancient retrotransposon was inserted into the genome of an eutherian ancestor and progressively integrated into host molecular networks while acquiring crucial functions for the organism. We suggest that this process of strict neofunctionalization was channeled by the inherited structural properties of the ancestral transposable element, facilitating the evolution of protein interactions. This domesticated gene subsequently duplicated, generating the *Bex/Tceal* cluster, whose expression is enriched in neural tissues. By generating new alleles for one of its members, *Bex3*, we show that this gene is involved in the development, maintenance, and function of specific areas of the central nervous system. The mutant alleles show subtle anatomical alterations in skull morphology, and important variations in cortical and subcortical neuron populations, as well as electrophysiological hippocampal imbalance, which can explain some of the behavioral changes identified. We also find a disruption in the mTOR pathway that might explain the molecular cause underlying some of the observed phenotypes. It remains to be elucidated if other genes in the cluster, most of which are highly expressed in the central nervous system, play similar or complementary roles in the establishment of higher neurological functions in placental mammals, or into which existing signaling pathways have been incorporated.

## Methods

### Study design

Our original hypothesis posited that recently improved genome and repetitive element annotations may help uncover new events of transposon domestication in the human and mouse genomes. Thus, the primary objective was to identify previously undetected TE-derived genes in the human and mouse genomes. A bioinformatic genome-wide screening provided a list of putative candidates, including previously uncharacterized cases. We decided to study in depth the event that gave rise to the *Bex/Tceal* gene family, which originated at the base of the eutherian lineage. We hypothesized that they could have provided an evolutionary novelty and sought to functionally characterize one of its members, *Bex3*, to explore this possibility. Using the CRISPR/Cas9 technology in mice, we generated two mutant lines for this gene, namely $Bex3^{KO}$ and $Bex3^{\Delta(24-72)}$, which were maintained in a CBA/C57Bl6 hybrid background. These mice, together with wild-type counterparts, were subjected to a battery of morphological, behavioral, physiological, histological, and molecular analyses. Additionally, chick embryos were employed as a non-eutherian animal model. All experiments were performed using 4 to 7 animals per experimental group, as presented in the manuscript. They were randomly chosen from the available animals in each case. For quantitative data, most

measurements were made automatically and, thus, not subject to operator bias. Additionally, the limited number of researchers working on each experiment prevented blinding procedures. No data were excluded from the analysis.

### Identification of domesticated transposon candidates

Gene and repetitive element annotations for the hg38 human and mm10 mouse genome assemblies were downloaded from the table browser tool of the UCSC website, selecting UCSC genes and RepeatMasker tracks, respectively. In order to discard non-TE elements, we filtered the RepeatMasker output as in [93]. These gene and repetitive element annotations were directly compared, and the results were filtered according to the following criteria: (i) overlap above 50% across the candidate gene coding region; (ii) txCdsPredict score above 800, which is approximately 90% predictive of protein-coding genes (it considers the length of the ORF, the presence of a Kozak consensus sequence, nonsense mediated decay mechanisms, and upstream ORFs); and (iii) conservation of the ORF in more than one species. Mammalian species with genome available at Ensembl genome browser (http://www.ensembl.org) were considered when checking for ORF conservation.

### Sequence retrieval

BLASTn and BLAT searches in NCBI and UCSC databases using human sequences were carried out in order to identify *Bex/Tceal* family genes in eutherian species. To clarify the orthology between genes undergoing gene conversion, surrounding intergenic sequences were added to the corresponding human query. The genomic assemblies used were as follows: GRCh38/hg38 for human (*Homo sapiens*); Broad CanFam3.1/canFam3 for dog (*Canis familiaris*); GRCm38/mm10 for mouse (*Mus musculus*); Bos_taurus_UMD_3.1.1/bosTau8 for cow (*Bos taurus*); Broad/equCab2 for horse (*Equus caballus*); Baylor/dasNov3 for nine-banded armadillo (*Dasypus novemcinctus*); Broad/choHof1 for Hoffmann's two-toed sloth (*Choloepus hoffmanni*); Broad v1.0/triMan1 for Florida manatee (*Trichechus manatus latirostris*); Broad/monDom5 for gray short-tailed opossum (*Monodelphis domestica*); WTSI Devil_ref v7.0/sarHar1 for Tasmanian devil (*Sarcophilus harrisii*), TWGS Meug_1.1/macEug2 for tammar wallaby (*Macropus eugenii*). The consensus sequences of TE subfamilies were retrieved from the RepBase-derived RepeatMasker library update 20170127 (http://www.girinst.org/server/RepBase/).

### Evolutionary analyses and secondary structure prediction

Nucleotide and protein sequences were aligned using the L-INS-i iterative refinement method of MAFFT [94], and the resulting alignments were edited with Jalview [95]. The phylogenetic reconstruction was performed using IQ-TREE [96] and built-in software ModelFinder [97]. Branch support was calculated running 1000 replicates of the SH-like approximate likelihood ratio test [98] and ultrafast bootstrap [99]. Phylogenetic trees were visualized and edited with FigTree v1.4.2 [100] and Dendroscope 3 [101]. PCOILS was used for coiled coil prediction [102]. PSI-PRED 4.0 [103, 104] and DISOPRED3 [105] were used for $\alpha$-helix and protein disorder prediction, respectively. Finally, the positive selection analysis was performed using HyPhy [106], specifically the

MEME [107] and aBSREL [108] methods. Taking into account the divergence observed between the *Bex* and *Tceal* subfamilies, the analysis was split in two parts to avoid saturation of substitutions.

### Differential expression analysis of eutherian adult tissues

RNA-seq data from eight homologous adult organs of dog, cow, and nine-banded armadillo were downloaded from SRA database: SRP016501 (GSE41637, *Bos taurus*), SRP114662 (GSE20113, *Canis familiaris*), and SRP012922 (GSE106077, *Dasypus novemcinctus*). Protein-coding cDNA sequences were downloaded from Ensembl (http://www.ensembl.org) as a transcriptomic index to map against. When *Bex/Tceal* orthologs were found to be incomplete or absent in these files, we replaced the partial sequences or introduced the missing ones, respectively. RNA-seq data was trimmed to 50 base pairs and mapped using Bowtie (allowing no more than two mismatches and discarding reads that mapped more than once) against their respective libraries for each species, and expression levels were calculated correcting for the effective length of each transcript (read-long positions repeated in other transcripts were excluded) to obtain cRPKM metrics [109]. cRPKM values for human and mouse were obtained from VAST DB [110]. Heatmaps were produced with Heatmapper [111] using Pearson's distance measurement and average linkage as the clustering method.

### Expression of *BEX/TCEAL* genes in autism spectrum disorder and schizophrenia

Differential expression of *BEX/TCEAL* genes was assessed using transcriptomic data from all publicly available human transcriptomic datasets for schizophrenia and autism spectrum disorder, either in GEO (http://www.ncbi.nlm.nih.gov/geo) or published articles (Additional file 1: Table S2). Statistical analysis of these data is specified in the corresponding "Statistical analysis" section.

### In situ hybridization

Mouse embryos were collected at E13.5, cryoprotected, and cut in 20-μm-thick sections. Primer pairs (Additional file 1: Table S4) were designed in order to amplify murine *Bex/Tceal* genes by PCR using cDNA. Next, digoxigenin-labeled RNA antisense probes were synthesized, and in situ hybridization was performed as described elsewhere [112]. Images were obtained with a Leica MZ16 F stereomicroscope.

### Generation of transgenic mice using the CRISPR-Cas9 system

Two CRISPR guide RNAs (sgRNAs) targeting the mouse *Bex3* gene were designed on exon 1 using CRISPR DESIGN (http://crispr.mit.edu) and CRISPRSCAN [113], considering potential off-target effects and predicted functional activity. They were generated by in vitro transcription from sgRNA DNA templates as previously described [114]. Briefly, for each sgRNA, two complementary oligos containing the CRISPR-Cas9 target sequence were annealed leaving BbsI-compatible overhangs. Oligo sequences for each sgRNA can be found in Additional file 1: Table S4. Annealed products were then cloned into the pgRNAbasic plasmid, a kind gift from Dr. Moises Mallo (Instituto Gulbenkian de Ciencia, Portugal), at the *Bbs*I site located downstream the T7 promoter and upstream the universal tracrRNA sequence necessary for sgRNA folding and

activity. Plasmids were linearized with *Fsp*I and the purified product used for T7 in vitro transcription according to the manufacturer's instructions (New England Biolabs). sgRNAs were purified by phenol:chloroform extraction after DNaseI treatment (Roche). Next, in vitro digestion assays were performed to evaluate sgRNA activity [115]. In brief, a DNA template containing the CRISPR-Cas9 target sites was amplified by nested PCR with the primers listed in Additional file 1: Table S4. Purified PCR products (200 ng) were incubated for 3 h at 37 °C with Cas9 protein (20 ng/μl, Addgene vector #47327) and sgRNA (2.5 ng/μl) in digestion buffer (20 mM HEPES pH 7.5, 150 mM KCl, 0.5 mM DTT, 0.1 mM EDTA, 10 mM $MgCl_2$). Cleavage efficiency was then analyzed by electrophoresis in a 1% agarose gel stained with ethidium bromide.

Cas9 mRNA (100 ng/μl, SBI) and sgRNA (10 ng/μl) were co-injected into the cytoplasm of CBA/C57Bl6 fertilized eggs using standard methods. Deletions in $F_0$ pups were detected by nested PCR (see Additional file 1: Table S4) of genomic DNA obtained from tail biopsies, and confirmed by sequencing after TA cloning into pCR2.1 (Invitrogen). Mutant $F_0$ carriers were crossed with wild-type CBA/C57Bl6 hybrids, and their descendants were likewise analyzed to identify the specific deletion allele transmitted. Mutant lines with deletions in mm10 chrX:136271359-136271505, for $Bex3^{\Delta(24-72)}$, and in chrX:136271327-136271522, for $Bex3^{KO}$, were established and maintained in a hybrid background.

### RT-PCR analyses

Brain RNA was extracted from three adult male wild-type and mutant mice using Tri-Pure isolation reagent (Roche), and cDNA was then synthesized with the first-strand cDNA synthesis kit for RT-PCR (AMV) (Roche), according to the manufacturer's instructions. *Bex3* and *actin* expression were analyzed with the primers listed in Additional file 1: Table S4.

### Morphometric analysis of mouse skulls

Control and mutant adult male mice at 6 to 8 weeks of age were sacrificed by cervical dislocation. Their heads were severed, skinned, defleshed, and incubated for 2–3 days in 2% NaOH/PBS in agitation to remove all remaining soft tissue. Undigested tissues where removed with forceps and skulls washed 3 times in PBS for 20 min. Samples were oriented and photographed using a high-resolution 3-CCD JVC camera (model KY-F55B) fitted onto a Nikon SMZ1500 stereomicroscope. Images were assembled, and 25 anatomical measurements were taken in Adobe Photoshop. Results were normalized to maximum skull length (D0 in Fig. 3) and relativized with respect to controls.

### Histological staining

Adult control and *Bex3* mutant mice were anesthetized and transcardially perfused for 20 min with 4% PFA/PBS. Brains were removed, post-fixed 24–48 h in 4% PFA/PBS, cryoprotected with 30% sucrose/PBS overnight, frozen in isopentane, and stored at −80 °C until use. Brains were sectioned coronally (30 μm), sections were collected in cryoprotectant solution (85% glycerol, 100% ethylene glycol, 0.1 M PBS) and kept at −20 °C until analysis. In the case of Nissl staining, sections were dehydrated and

mounted (Eukitt). Slide images were captured using a Nanozoomer slide scanner (Hamamatsu) and analyzed by measuring the maximum width (mm) and height (mm) in 5 sections per animal, corresponding to different anterior-posterior levels based on Bregma. Measurements were performed by using NanoZoomer Digital Pathology software. Ventricular surface was measured in 2 histological sections per animal by using ImageJ software. Additionally, immunohistochemistry (IHC) labelling was performed on 50-μm coronal brain cryosections with the following primary antibodies: mouse anti-NeuN (1:3000, Millipore), rabbit anti-parvalbumin (1:3000, Swant), mouse anti-calretinin (1:1000, Novocastra), rat anti-GFAP (1:3000, Calbiochem), and goat anti-doublecortin (1:500, Santa Cruz). Immunoreactivity was developed with DAB-peroxidase reaction. To minimize variability, at least 2–3 sections from each area were analyzed per animal on a bright-field DMRB RFY HC microscope (Leica). In each section, the total number of labeled cells per area of tissue was quantified using ImageJ software.

Chick embryos were fixed for 2–4 h at 4 °C in 4% (w/v) paraformaldehyde in PB, washed in PBS, and vibratome sectioned (45 μm). BrdU detection and immunostaining were performed following standard procedures [116, 117]. The following antibodies were used: rabbit anti-GFP (1:500, Invitrogen), rat anti-BrdU (1:500, AbDSerotec), rabbit anti-Sox2 (1:500, Invitrogen), Alexa488- and Alexa562-conjugated antibodies (Invitrogen). The sections were recorded using a Leica SPE confocal microscope. Cell counting was carried out on 10–17 pictures obtained from 5 to 7 chick embryos per experimental condition.

**Behavioral tests**

Behavioral tests were performed in 3- to 5-month-old mice in a room with constant sound and light after 1 h of habituation. In all tests, during mice manipulation and behavioral evaluation, the researcher was blind to mice genotypes. The order in which the tests were run was always the same: open field, three-chambered test, nesting test, startle, Y-maze, object recognition memory, and step-down passive avoidance.

*Open field test*

Motor activity was assessed in an open field for 5 min (38 × 21 × 15 cm: Cybertec S.A., Madrid, Spain) as previously described [92]. This apparatus, which is coupled with infrared (IR) emitters and sensors connected to a computer, allowed to quantify the number of times mice interrupted the IR beams/min. In parallel, sessions were video recorded in order to evaluate rearing and grooming.

*Three-chambered sociability test*

Sociability was analyzed in an arena (60 × 60 cm) divided in three equal compartments. Each lateral compartment contained a clear Plexiglas cylinder (each 7 cm in diameter, 12 cm tall) with multiple holes (1 cm in diameter) to allow auditory, visual, and olfactory interaction between the stimulus mouse and the test mouse placed inside and outside of the cylinder, respectively. The paradigm, lasting 5 min per animal, consisted in a three-stage procedure: During the *habituation phase,* the test mouse was allowed to explore the

apparatus with the cylinders empty. In the *social familiar subject phase*, a conspecific from its same home cage was placed in one of the cylinders while the other remained empty. Finally, the *social stranger subject phase* was performed with an unfamiliar age- and sex-matched CD1 mouse placed in one of the cylinders maintaining the other empty. Social interaction was evaluated by the time that test mice spent sniffing each cylinder.

### Nesting test
The ability to build a nest was assessed following previously described procedures [118].

### Acoustic startle reflex and prepulse inhibition
During training, the mouse was placed in the startle chamber (Cibertec S.A., Madrid, Spain) for 3 min for acclimation; then, baseline startle responses were measured and averaged from 25 to 30 recordings after the presentation of 20 sounds (125 dB, 100 ms long). From this phase, the average response and peak latencies, as well as the peak response, were determined. During prepulse inhibition (PPI) trials, the same sound was preceded (250 ms) by a prepulse stimulus of 85 dB, 50 ms long. Trials including prepulse stimuli were randomly presented with normal startle stimuli, the final total being 25 of each, and the proportion of PPI was determined as $[(1 - \text{prepulse/startle}) \times 100]$. The ambient background noise was 70 dB.

### Y-maze
Each mouse was placed in the center of a Y-maze consisting of three equally sized arms ($8 \times 40 \times 20$ cm), with white opaque walls at a 120° angle from each other, and allowed to freely explore the arms during 5 min. The number of arm entries and triads were used to calculate the alternation index. An entry was considered to occur when all four limbs were within the arm.

### Object recognition memory
The object recognition protocol was described extensively elsewhere [92]. Briefly, two equal objects were placed in a rectangular arena ($55 \times 40 \times 40$ cm) during the training phase. The next day, one object was replaced by a novel one, and the animal's memory of the original object was assessed by comparing the amount of time spent actively exploring the novel object against that for the familiar one using a discrimination index $[\text{DI} = (t_{\text{novel}} - t_{\text{familiar}})/(t_{\text{novel}} + t_{\text{familiar}})$. Exploration of an object was defined as directing the nose toward the object at a distance of $\leq 1.5$ cm or touching the object with the nose or vibrissae. Circling or sitting on the object were not considered exploratory behaviors.

### Step-through passive avoidance test
During habituation, mice were allowed to explore for 1 min a chamber ($47 \times 18 \times 26$ cm, Ugo Basile) symmetrically divided into one light and one dark compartment separated by a door. During the training phase, mice were confined to the light compartment for 30 s and then allowed to access the dark compartment. Once inside, the door was closed automatically and the mice received an electrical stimulation (0.3 mA, 5 s) through the metal floor. Retention tests were performed the next day. Here, the latency

to enter into the dark compartment (escape latency) was assessed as a measure of memory retention.

### Electrophysiological recordings

Hippocampal slices were prepared as previously described [119, 120]. Five-month-old mice were anesthetized with isoflurane (2%) and decapitated for slice preparation. Briefly, after decapitation, the whole brain containing the two hippocampi, was placed into ice-cold solution (I) consisting of the following (in mM): 126 NaCl, 3 KCl, 1.25 $KH_2PO_4$, 2 $MgSO_4$, 2 $CaCl_2$, 26 $NaHCO_3$, and 10 glucose (pH 7.2, 300 mOsm), positioned on the stage of a vibratome slicer and cut to obtain transverse hippocampal slices (350 μm). Slices were maintained continuously oxygenated for at least 1 h before use. All experiments were carried out at room temperature (22–25 °C). For experiments, slices were continuously perfused with the solution described above. Whole-cell patch clamp recording of pyramidal cells located in the CA2 field of the hippocampus was obtained under visual guidance by IR differential interference contrast (DIC) microscopy and were verified as pyramidal cells through their characteristic voltage response to a current step protocol. Neurons were recorded in the current-clamp configuration with a Multiclamp 700B patch clamp amplifier. Data were acquired using pCLAMP 10.2 software (Molecular Devices). To record evoked excitatory postsynaptic currents (eEPSCs), electrical pulses were delivered to Schaffer collateral axons. To evoke inhibitory postsynaptic currents (eIPSCs), electrical pulses were delivered to interneurons situated in the *stratum oriens*. Spontaneous synaptic activity (sPSCs), either excitatory (sEPSCs) or inhibitory (sIPSCs), was recorded from hippocampal CA2 neurons.

Patch electrodes were pulled from borosilicate glass and had a resistance of 4–7 MΩ when filled with the following (in mM): 120 CsCl, 8 NaCl, 1 $MgCl_2$, 0.2 $CaCl_2$, 10 HEPES, 2 EGTA, and 20 QX-314 (pH 7.2, 290 mOsm). Experiments were performed at − 70 mV. Only cells with a stable resting membrane potential of − 55 mV were assessed, and the cell recordings were discarded if the series resistance changed by more than 15%. All recordings were low-pass filtered at 2 kHz and acquired at 10 kHz. Excitatory postsynaptic currents (evoked, eEPSCs, and spontaneous, sEPSCs) were isolated by adding bicuculline (20 μM) to the perfusion solution to block GABAA receptors. Inhibitory postsynaptic currents (evoked, eIPSCs, and spontaneous, sIPSCs) were isolated adding D-AP5 (50 μM) and NBQX (10 μM) to the perfusion solution to block NMDA receptor and AMPA/Kainate receptor-mediated currents, respectively. Signals were averaged every 12 traces. Spontaneous recordings consisted of 60 sweeps/5 s long that were analyzed for amplitude and frequency of detected events.

### Western blotting

Total brain tissue was homogenized in RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.1% (w/v) SDS), containing protease inhibitor cocktail (Roche) and phosphatase inhibitors (2 mM sodium orthovanadate, 1 mM sodium pyrophosphate, 10 mM sodium fluoride). Protein concentration was measured using the bicinchoninic acid (BCA) protein assay as specified by the manufacturer (Pierce, Thermo Fisher Scientific). Samples were resolved by SDS-polyacrylamide

gels and transferred onto nitrocellulose membranes. Blocking and incubation with primary and secondary antibodies were performed following the manufacturer instructions. Membranes were developed with the ECL system (GE Healthcare), and acquired images were quantified using ImageJ software. Tubulin or GAPDH loading controls were used when needed. The primary antibodies used were the following: S6 1/1000 (#2217; Cell Signaling), p-S6 (Ser235/236) 1/1000 (#2217; Cell Signaling), S6K1 1/1000 (#9202; Cell Signaling), p-S6K1 (Thr389) 1/1000 (#9205; Cell Signaling), 4E-BP1 1/1000 (#9644; Cell Signaling), AKT 1/1000 (C-20, Santa Cruz Biotechnology), p-AKT (Ser473) 1/1000 (#9271; Cell Signaling), PKCα 1/1000 (10/2018, Cell Signaling), PKCγ 1/1000 (C-4, SC-166385, Santa Cruz Biotechnology), NDRG1 1/1000 (B-5, SC-398291, Santa Cruz Biotechnology), p-NDRG1 (Thr346) 1/1000 (#3217; Cell Signaling), GAPDH (6C5, ab 8245, Abcam), and tubulin 1/1000 (#9026; Sigma-Aldrich). The secondary antibodies used were HRP-labeled anti-mouse 1/2000 (P447-01, Vector) and anti-rabbit 1/2000 (P217-02, Vector).

### HALEX ancestral protogene reconstruction and in ovo electroporation

In order to obtain an approximate reconstruction of the ORF of the ancestral *HALEX* protogene, a segment of the consensus sequences of HAL1b and inverted L1MEe derived from RepeatMasker, together with the sequence of the *TCEAL7* human gene corresponding to the hypothetical target site duplication of an ancestral L1ME element, were combined (Additional file 1: Table S4). Finally, the synthesis of an artificial gene cloned into a pcDNA™3.1/*myc*-His was ordered to GenScript.

Eggs from white-Leghorn chickens were incubated at 38.5 °C in an atmosphere with 70% humidity and staged according to the method of Hamburger and Hamilton [121]. In ovo electroporation was performed at stage HH11-12 (48 h of incubation) with DNA plasmids as described previously [117, 122]. Bromo-deoxyuridine (1 mM; Sigma) was injected into the lumen of the chick neural tube at 20 min before harvesting, to label dividing cells. The embryos were recovered at 24 h post-electroporation.

### Statistical analysis

Enrichment analyses for *BEX* and *TCEAL* genes in gene expression datasets for autism spectrum disorder and schizophrenia were performed for each dataset using a hypergeometric test; $p < 0.05$ was considered as a threshold for statistical significance. Other statistical analyses were performed using Student's two-tailed $t$ tests or one-way ANOVA with Tukey's multiple comparison tests, where appropriate. When data were not normally distributed, non-parametric Mann–Whitney tests were used to determine the statistical significance. Calculations were performed with GraphPad Prism version 6 or in R; $*p < 0.05$, $**p < 0.01$, $***p < 0.005$, $****p < 0.001$.

### Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13059-020-02172-3.

**Additional file 1:** Fig. S1 The *Tceal7* gene is derived from the domestication of L1 retrotransposon fragments. Fig. S2 The *Bex/Tceal* gene cluster was established before the diversification of extant eutherians. Fig. S3 Highly diverged BLX/TCLAL proteins share a coiled coil domain. Fig. S4 BLX/TCLAL proteins might have inherited some of their structural properties from the ancestral transposon. Fig. S5 Selection pressure analyses reveal signatures of positive selection in the *Bex/Tceal* genes. Fig. S6 A BGW-like sequence was already present in the G/A promoter of the last therian common ancestor. Fig. S7 *Bex/Tceal* genes show tissue-enriched expression patterns during

development. Fig. S8 *Bex3* and *Tceal7* genes, but not the ancestral HALEX element, induce cell proliferation in chicken neural tube. Fig. S9 The deletions introduced using CRISPR-Cas9 technology can be observed in the mRNA expressed from the *Bex3* mutant alleles. Fig. S10 CRISPR-Cas9-generated *Bex3* mutant alleles show subtle skull abnormalities. Fig. S11 *Bex3* mutant mice show normal acoustic startle reflex. Fig. S12 *Bex3* deficiency leads to aberrant mTOR signaling in the brain. Table S1 Coding genes putatively derived from transposable elements in the human and mouse genomes. Table S2 Altered expression of *BEX* and *TCEAL* genes in subjects with autism spectrum disorder or schizophrenia. Table S3 Enrichment of differential gene expression in *BEX* and *TCEAL* gene families in subjects with autism spectrum disorder or schizophrenia. Table S4 Primers and reconstructed gene sequences used in this work. Supplementary references.

**Additional file 2.** Review history.

### Authors' contributions
F.N.P. designed the study, performed molecular biology and in situ hybridization experiments, did bioinformatics and histologic analyses, and wrote the manuscript. C.V.G. designed the study, performed molecular biology experiments, and wrote the manuscript. S.M. did molecular biology experiments, performed histological analyses, and wrote the manuscript. D.B. designed the study, performed bioinformatics analyses and molecular biology experiments, and wrote the manuscript. N.F.C. did bioinformatics analyses. J.L.F. supervised in situ hybridization experiments in mice. M.L.M. did all skull measurements and preliminary behavioral tests. M.A.N. performed behavioral tests. I.S.-P. performed behavioral tests. E.A.G. did bioinformatics analyses. F.U. supervised molecular biology analyses. C.H.U. conducted selection pressure analysis. P.C. performed bioinformatics analyses. R.F.M. performed excitation/inhibition balance experiments. A.R.M. supervised excitation/inhibition balance experiments. S.D'A. helped in writing the manuscript. B.C. supervised bioinformatics analyses and provided resources. G.M. supervised molecular biology experiments and provided resources. E.S. supervised molecular biology and histology experiments and provided resources. A.M.C. performed and supervised behavioral and histology experiments and provided resources. J.J.C. designed the study, generated the mouse mutant lines, performed and supervised molecular biology experiments, and provided resources. J.G.F. designed the study and provided resources and supervision. The author(s) read and approved the final manuscript.

### Authors' information
Twitter handles: @cvicgar (Cristina Vicente-García); @MirraSerena (Serena Mirra); @Noeliafdez6 (Noèlia Fernàndez-Castillo); @jlferran (José Luis Ferrán); @maclopmay (Macarena López-Mayorga); @yomellamocar (Carlos Herrera-Úbeda); @Salv_DANIELLO (Salvatore D'Aniello); @bcormand (Bru Cormand); @GMarfanyN (Gemma Marfany); @jcarvajal (Jaime J. Carvajal); @jordigarciafdez (Jordi Garcia-Fernàndez).

### Availability of data and materials
All data generated during this study are included in this published article and its supplementary information files. All data reanalyzed in this study is publicly available from the sources specified in the methods. Specifically, RNA-seq datasets SRP016501 (GSL41637, *Bos taurus*), SRP114662 (GSL20113, *Canis familiaris*), and SRP012922 (GSL106077, *Dasypus novemcinctus*) were downloaded from SRA database, while protein-coding cDNA sequences were obtained from Ensembl (http://www.ensembl.org), and cRPKM values for human and mouse tissues, from VASTDB [110]. Further, human transcriptomic datasets for schizophrenia and autism spectrum disorder were obtained from GEO database (GSE38322, GSE35978, GSE53987, GSE87610) or published articles (Additional file 1: Supplementary references [16, 18, 21, 26]).

### Ethics approval and consent to participate

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

[1]Department of Genetics, Microbiology and Statistics, Faculty of Biology, and Institut de Biomedicina (IBUB), University of Barcelona, 08028 Barcelona, Spain. [2]Centro Andaluz de Biología del Desarrollo, CSIC-UPO-JA, Universidad Pablo de Olavide, 41013 Sevilla, Spain. [3]Department of Cell Biology, Physiology and Immunology, and Institute of Neurosciences, University of Barcelona, 08028 Barcelona, Spain. [4]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [5]Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Instituto de Salud Carlos III (ISCIII), 28029 Madrid, Spain. [6]Department of Zoology, Charles University, Vinicna 7, 12844 Prague, Czech Republic. [7]Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, 08950 Barcelona, Spain. [8]Department of Human Anatomy, School of Medicine, University of Murcia and IMIB-Arrixaca Institute, 30120 Murcia, Spain. [9]Department of Physiology, Anatomy and Cell Biology, Universidad Pablo de Olavide, 41013 Sevilla, Spain. [10]Present Address: Instituto de Neurociencias de Alicante (Universidad Miguel Hernández - Consejo Superior de Investigaciones Científicas), Alicante, Spain. [11]Present Address: Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Neuropsychopharmacology and psychobiology research group, UCA, INiBICA, Cádiz, Spain. [12]Genome Architecture, Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), the Barcelona Institute of Science and Technology, 08003 Barcelona, Spain. [13]Universitat Pompeu Fabra (UPF), 08003 Barcelona, Spain. [14]Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, 80121 Naples, Italy. [15]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain.

### References

1. Tautz D, Domazet-Lošo T. The evolutionary origin of orphan genes. Nat Rev Genet. 2011;12:692–702.
2. Miller WJ, McDonald JF, Pinsker W. Molecular domestication of mobile elements. Genetica. 1997;100:261–70.
3. Fugmann SD. The origins of the Rag genes-from transposition to V(D) J recombination. Semin Immunol. 2010;22:10–6.
4. Kim AJ, Lee CS, Schlessinger D. Bex3 associates with replicating mitochondria and is involved in possible growth control of F9 teratocarcinoma cells. Gene. 2004;343:79–89.
5. Chien J, Narita K, Rattan R, Giri S, Shridhar R, Staub J, et al. A role for candidate tumor-suppressor gene TCEAL7 in the regulation of c-Myc activity, cyclin D1 levels and cellular transformation. Oncogene. 2008;27(58):7223–34.
6. Rattan R, Narita K, Chien J, Maguire JL, Shridhar S, Giri S. et al. TCEAL7, a putative tumor suppressor gene, negatively regulates NF-B pathway. Oncogene. 2009;29:1362–73.
7. Kazi JU, Kabir NN, Ronnstrand L. Brain-expressed X-linked (BEX) proteins in human cancers. Biochim Biophys Acta. 2015; 1856:226–33.
8. Gao W, Li JZ, Chen SQ, Chu CY, Chan JY, Wong TS. BEX3 contributes to cisplatin chemoresistance in nasopharyngeal carcinoma. Cancer Med. 2017;6(2):439–51.
9. Ward C, Cauchy P, Garcia P, Frampton J, Esteban MA, Volpe G. High WBP5 expression correlates with elevation of HOX genes levels and is associated with inferior survival in patients with acute myeloid leukaemia. Sci Rep. 2020;10(1):3505.
10. Schwarz BA, Cetinbas M, Clement K, Walsh RM, Cheloufi S, Gu H, et al. Prospective isolation of poised iPSC intermediates reveals principles of cellular reprogramming. Cell Stem Cell. 2018;23(2):289–305.
11. Calvo L, Anta B, López-Benito S, Martin-Rodriguez C, Lee FS, Pérez P, et al. Bex3 dimerization regulates NGF-dependent neuronal survival and differentiation by enhancing trkA gene transcription. J Neurosci. 2015;35:7190–202.
12. Vilar M, Murillo-Carretero M, Mira H, Magnusson K, Besset V, Ibáñez CF. Bex1, a novel interactor of the p75 neurotrophin receptor, links neurotrophin signaling to the cell cycle. EMBO J. 2006;25:1219–30.
13. Judd J, Lovas J, Huang GN. Defined factors to reactivate cell cycle activity in adult mouse cardiomyocytes. Sci Rep. 2019; 9(1):18830.
14. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature. 2000;403:785–9.
15. Blaise S, de Parseval N, Bénit L, Heidmann T. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc Natl Acad Sci U S A. 2003;100:13013–8.
16. Robertson HM, Zumpano KL. Molecular evolution of an ancient *mariner* transposon, *Hsmar1*, in the human genome. Gene. 1997;205:203–17.
17. Hayward A, Ghazal A, Andersson G, Andersson L, Jern P. ZBED evolution: repeated utilization of DNA transposons as regulators of diverse host functions. PLoS One. 2013;8:e59940.
18. Campillos M, Doerks T, Shah PK, Bork P. Computational characterization of multiple gag-like human proteins. Trends Genet. 2006;22:585–9.
19. Ostertag EM, Kazazian J. Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome Res. 2001;11:2059–65.

20. Athanikar JN, Badge RM, Moran JV. A YY1-binding site is required for accurate human LINE-1 transcription initiation. Nucleic Acids Res. 2004;32:3846–55.
21. Winter EE, Ponting CP. Mammalian BEX, WEX and GASP genes: coding and non-coding chimaerism sustained by gene conversion events. BMC Evol Biol. 2005;5:54.
22. Giordano J, Ge Y, Gelfand Y, Abrusán G, Benson G, Warburton PE. Evolutionary history of mammalian transposons determined by genome-wide defragmentation. PLoS Comput Biol. 2007;3:1321–34.
23. Cabral KMS, Raymundo DP, Silva VS, Sampaio LAG, Johanson L, Hill LF, et al. Biophysical studies on BEX3, the p75NTR-associated cell death executor, Reveal a High-Order Oligomer with Partially Folded Regions. PLoS One. 2015;10: e0137916.
24. Fernandez EM, Díaz-Ceso MD, Vilar M. Brain expressed and X-linked (Bex) proteins are intrinsically disordered proteins (IDPs) and form new signaling hubs. PLoS One. 2015;10:e0117206.
25. do Amaral MJ, Araujo TS, Díaz NC, Accornero F, Polycarpo CR, Cordeiro Y, et al. Phase separation and disorder-to-order transition of human brain expressed X-Linked 3 (hBEX3) in the presence of small fragments of tRNA. J Mol Biol. 2020; 432(7):2319–48.
26. Zhang L. Adaptive evolution and frequent gene conversion in the brain expressed X-linked gene family in mammals. Biochem Genet. 2008;46(5–6):293–311.
27. Ratnakumar A, Mousset S, Glémin S, Berglund J, Galtier N, Duret L, et al. Detecting positive selection within genomes: the problem of biased gene conversion. Philos Trans R Soc B Biol Sci. 2010;365:2571–80.
28. López-Doménech G, Serrat R, Mirra S, D'Aniello S, Somorjai I, Abad A, et al. The Eutherian Armcx genes regulate mitochondrial trafficking in neurons and interact with Miro and Trak2. Nat Commun. 2012;3:814.
29. Shi X, Garry DJ. Myogenic regulatory factors transactivate the Tceal7 gene and modulate muscle differentiation. Biochem J. 2010;428:213–21.
30. Iseki H, Takeda A, Andoh T, Takahashi N, Kurochkin IV, Yarmishyn A, et al. Human Arm protein lost in epithelial cancers, on chromosome X 1 (ALEX1) gene is transcriptionally regulated by CREB and Wnt/beta-catenin signaling. Cancer Sci. 2010;101:1361–6.
31. Andrianaki A, Siapati LK, Hirata RK, Russell DW, Vassilopoulos G. Dual transgene expression by foamy virus vectors carrying an endogenous bidirectional promoter. Gene Ther. 2010;17:380–8.
32. Alvarez E, Zhou W, Witta SE, Freed CR. Characterization of the Bex gene family in humans, mice, and rats. Gene. 2005; 357:18–28.
33. Koo JH, Saraswati M, Margolis FL. Immunolocalization of Bex protein in the mouse brain and olfactory system. J Comp Neurol. 2005;487:1–14.
34. Ito K, Yamazaki S, Yamamoto R, Tajima Y, Yanagida A, Kobayashi T, et al. Gene targeting study reveals unexpected expression of brain-expressed X-linked 2 in endocrine and tissue stem/progenitor cells in mice. J Biol Chem. 2014;289: 29892–911.
35. Accornero F, Schips TG, Petrosino JM, Gu S-Q, Kanisicak O, van Berlo JH, et al. BEX1 is an RNA-dependent mediator of cardiomyopathy. Nat Commun. 2017;8:1875.
36. Yu W, Yaping L, Mingjun W, Jie H, Xiaogang L, Gang L. BEX4 upregulation alters Sertoli cell growth properties and protein expression profiles: An explanation for cadmium-induced testicular Sertoli cell injury. J Biochem Mol Toxicol. 2017;31 https://doi.org/10.1002/jbt.21908.
37. Gu Y, Wei W, Cheng Y, Wan B, Ding X, Wang H, et al. A pivotal role of BEX1 in liver progenitor cell expansion in mice. Stem Cell Res Ther. 2018;9:164.
38. Myers A, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, et al. A survey of genetic human cortical gene expression. Nat Genet. 2007;39(12):1494–9.
39. Khazaei MR, Halfter H, Karimzadeh F, Koo JH, Margolis FL, Young P. Bex1 is involved in the regeneration of axons after injury. J Neurochem. 2010;115:910–20.
40. Yamamoto T, Wilsdon A, Joss S, Isidor B, Erlandsson A, Suri M, et al. An emerging phenotype of Xq22 microdeletions in females with severe intellectual disability, hypotonia and behavioral abnormalities. J Hum Genet. 2014;59:300–6.
41. Shirai K, Higashi Y, Shimojima K, Yamamoto Y. An Xq22.1q22.2 nullisomy in a male patient with severe neurological impairment. Am J Med Genet Part A. 2017;173:1124–7.
42. Hijazi H, Coelho FS, Gonzaga-Jauregui C, Bernardini L, Mar SS, Manning MA, et al. Xq22 deletions and correlation with distinct neurological disease traits in females: further evidence for a contiguous gene syndrome. Hum Mutat. 2020;41: 150–68.
43. Mukai J, Hachiya T, Shoji-Hoshino S, Kimura MT, Nadano D, Suvanto P, et al. NADE, a p75NTR-associated cell death executor, is involved in signal transduction mediated by the common neurotrophin receptor p75NTR. J Biol Chem. 2000;275:17566–70.
44. Mukai J, Shoji S, Kimura MT, Okubo S, Sano H, Suvanto P, et al. Structure function analysis of NADE: identification of regions that mediate nerve growth factor-induced apoptosis. J Biol Chem. 2002;277:13973–82.
45. Fombonne E, Rogé B, Claverie J, Courty S, Frémolle J. Microcephaly and macrocephaly in autism. J Autism Dev Disord. 1999;29:113–9.
46. Wright IC, Rabe-Hesketh S, Woodruff PWR, David AS, Murray RM, Bullmore ET. Meta-analysis of regional brain volumes in schizophrenia. Am J Psychiatry. 2000;157:16–25.
47. Turner AH, Greenspan KS, van Erp TGM. Pallidum and lateral ventricle volume enlargement in autism spectrum disorder. Psychiatry Res Neuroimaging. 2016;252:40–5.
48. Lijam N, Paylor R, McDonald MP, Crawley JN, Deng CX, Herrup K, et al. Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. Cell. 1997;90:895–905.
49. El-Kordi A, Winkler D, Hammerschmidt K, Kästner A, Krueger D, Ronnenberg A, et al. Development of an autism severity score for mice using Nlgn4 null mutants as a construct-valid model of heritable monogenic autism. Behav Brain Res. 2013;251:41–9.
50. Kohl S, Heekeren K, Klosterkötter J, Kuhn J. Prepulse inhibition in psychiatric disorders--apart from schizophrenia. J Psychiatr Res. 2013;47:445–52.
51. Marín O. Interneuron dysfunction in psychiatric disorders. Nat Rev Neurosci. 2012;13:107–20.

52. Pelkey KA, Chittajallu R, Craig MT, Tricoire L, Wester JC, McBain CJ. Hippocampal gabaergic inhibitory interneurons. Physiol Rev. 2017;97:1619–747.
53. Hitti FL, Siegelbaum SA. The hippocampal CA2 region is essential for social memory. Nature. 2014;508:88–92.
54. Dudek SM, Alexander GM, Farris S. Rediscovering area CA2: unique properties and functions. Nat Rev Neurosci. 2016;17:89–102.
55. Yasui S, Tsuzaki K, Ninomiya H, Floricel F, Asano Y, Maki H, et al. The TSC1 gene product hamartin interacts with NADE. Mol Cell Neurosci. 2007;35:100–8.
56. Huang J, Dibble CC, Matsuzaki M, Manning BD. The TSC1-TSC2 complex is required for proper activation of mTOR complex 2. Mol Cell Biol. 2008;28:4104–15.
57. Lipton JO, Sahin M. The neurology of mTOR. Neuron. 2014;84:275–91.
58. McLaughlin RN, Young JM, Yang L, Neme R, Wichman HA, Malik HS. Positive selection and multiple losses of the LINE-1-derived L1TD1 gene in mammals suggest a dual role in genome defense and pluripotency. PLoS Genet. 2014;10:e1004531.
59. Shaheen M, Williamson E, Nickoloff J, Lee S-H, Hromas R. Metnase/SETMAR: a domesticated primate transposase that enhances DNA repair, replication, and decatenation. Genetica. 2010;138:559–66.
60. Gewurz BE, Towfic F, Mar JC, Shinners NP, Takasaki K, Zhao B, et al. Genome-wide siRNA screen for mediators of NF-κB activation. Proc Natl Acad Sci U S A. 2012;109:2467–72.
61. Arroyo R, Suñé G, Zanzoni A, Duran-Frigola M, Alcalde V, Stracker TH, et al. Systematic identification of molecular links between core and candidate genes in breast cancer. J Mol Biol. 2015;427:1436–50.
62. Hu Z, Wang Y, Huang F, Chen R, Li C, Wang F, et al. Brain-expressed X-linked 2 is pivotal for hyperactive mechanistic target of rapamycin (mTOR)-mediated tumorigenesis. J Biol Chem. 2015;290:25756–65.
63. Lee JK, Lee J, Go H, Lee CG, Kim S, Kim HS, et al. Oncogenic microtubule hyperacetylation through BEX4-mediated sirtuin 2 inhibition. Cell Death Dis. 2016;7:e2336.
64. Zhao Z, Li J, Tan F, Gao S, He J. mTOR up-regulation of BEX4 promotes lung adenocarcinoma cell proliferation by potentiating OCT4. Biochem Biophys Res Commun. 2018;500:302–9.
65. Haynes C, Oldfield CJ, Ji F, Klitgord N, Cusick ME, Radivojac P, et al. Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. PLoS Comput Biol. 2006;2(8):e100.
66. Laplante M, Sabatini DM. MTOR signaling in growth control and disease. Cell. 2012;149:274–93.
67. Huang J, Manning BD. A complex interplay between Akt, TSC2 and the two mTOR complexes. Biochem Soc Trans. 2009;37:217–22.
68. Martin TD, Dennis MD, Gordon BS, Kimball SR, Jefferson LS. mTORC1 and JNK coordinate phosphorylation of the p70S6K1 autoinhibitory domain in skeletal muscle following functional overloading. Am J Physiol Endocrinol Metab. 2014;306:L1397–405.
69. Biever A, Valjent E, Puighermanal E. Ribosomal protein S6 phosphorylation in the nervous system: from regulation to function. Front Mol Neurosci. 2015;8:75.
70. Kokošar J, Kordiš D. Genesis and regulatory wiring of retroelement-derived domesticated genes: a phylogenomic perspective. Mol Biol Evol. 2013;30:1015–31.
71. Faulkner GJ, Kimura Y, Daub CO, Wani S, Plessy C, Irvine KM, et al. The regulated retrotransposon transcriptome of mammalian cells. Nat Genet. 2009;41(5):563–71.
72. Villanueva-Cañas JL, Ruiz-Orera J, Agea MI, Gallo M, Andreu D, Albà MM. New genes and functional innovation in mammals. Genome Biol Evol. 2017;9:1886–900.
73. De Angelis MT, Russo F, D'Angelo F, Federico A, Gemei M, Del Vecchio L, et al. Novel pancreas organogenesis markers refine the pancreatic differentiation roadmap of embryonic stem cells. Stem Cell Rev Rep. 2014;10:269–79.
74. Gergics P, Christian HC, Choo MS, Ajmal A, Camper SA. Gene expression in mouse thyrotrope adenoma: transcription elongation factor stimulates proliferation. Endocrinology. 2016;157:3631–46.
75. Mathys H, Davila-Velderrain J, Peng Z, Gao F, Mohammadi S, Young JZ, et al. Single-cell transcriptomic analysis of Alzheimer's disease. Nature. 2019;570:332–7.
76. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. Am J Hum Genet. 2009;84:524–33.
77. Silverman JL, Yang M, Lord C, Crawley JN. Behavioural phenotyping assays for mouse models of autism. Nat Rev Neurosci. 2010;11:490–502.
78. Sandi C, Haller J. Stress and the social brain: behavioural effects and neurobiological mechanisms. Nat Rev Neurosci. 2015;16:290–304.
79. del Pino I, Rico B, Marín O. Neural circuit dysfunction in mouse models of neurodevelopmental disorders. Curr Opin Neurobiol. 2018;48:174–82.
80. Langen M, Durston S, Kas MJH, van Engeland H, Staal WG. The neurobiology of repetitive behavior: … and men. Neurosci Biobehav Rev. 2011;35:356–65.
81. Selby L, Zhang C, Sun Q-Q. Major defects in neocortical GABAergic inhibitory circuits in mice lacking the fragile X mental retardation protein. Neurosci Lett. 2007;412:227–32.
82. Gant JC, Thibault O, Blalock LM, Yang J, Bachstetter A, Kotick J, et al. Decreased number of interneurons and increased seizures in neuropilin 2 deficient mice: implications for autism and epilepsy. Epilepsia. 2009;50:629–45.
83. Gogolla N, LeBlanc JJ, Quast KB, Südhof TC, Fagiolini M, Hensch TK. Common circuit defect of excitatory-inhibitory balance in mouse models of autism. J Neurodev Disord. 2009;1:172–81.
84. Lawrence YA, Kemper TL, Bauman ML, Blatt GJ. Parvalbumin-, calbindin-, and calretinin-immunoreactive hippocampal interneuron density in autism. Acta Neurol Scand. 2010;121:99–108.
85. Zikopoulos B, Barbas H. Altered neural connectivity in excitatory and inhibitory cortical circuits in autism. Front Hum Neurosci. 2013;7:609.
86. Stoner R, Chow ML, Boyle MP, Sunkin SM, Mouton PR, Roy S, et al. Patches of disorganization in the neocortex of children with autism. N Engl J Med. 2014;370:1209–19.
87. Wöhr M, Orduz D, Gregory P, Moreno H, Khan U, Vörckel KJ, et al. Lack of parvalbumin in mice leads to behavioral deficits relevant to all human autism core symptoms and related neural morphofunctional abnormalities. Transl Psychiatry. 2015;5:e525.

88.  Cai Y, Tang X, Chen X, Li X, Wang Y, Bao X, et al. Liver X receptor β regulates the development of the dentate gyrus and autistic-like behavior in the mouse. Proc Natl Acad Sci U S A. 2018;115:L2725–33.
89.  Goh S, Peterson BS. Imaging evidence for disturbances in multiple learning and memory systems in persons with autism spectrum disorders. Dev Med Child Neurol. 2012;54:208–13.
90.  Baarendse PJJ, Van Grootheest G, Jansen RF, Pieneman AW, Ögren SO, Verhage M, et al. Differential involvement of the dorsal hippocampus in passive avoidance in C57Bl/6J and DBA/2J mice. Hippocampus. 2008;18:11–9.
91.  Winters BD, Saksida LM, Bussey TJ. Object recognition memory: neurobiological mechanisms of encoding, consolidation and retrieval. Neurosci Biobehav Rev. 2008;32:1055–70.
92.  Suárez-Pereira I, Canals S, Carrión AM. Adult newborn neurons are involved in learning acquisition and long-term memory formation: the distinct demands on temporal neurogenesis of different cognitive tasks. Hippocampus. 2015;25:51–61.
93.  Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay LA, Bourque G, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. PLoS Genet. 2013;9:e1003470.
94.  Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
95.  Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25:1189–91.
96.  Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.
97.  Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9.
98.  Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59:307–21.
99.  Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Molecular biology and evolution. Mol Biol Evol. 2018;35:518–22.
100. Rambaut A. FigTree v1.4.2, a graphical viewer of phylogenetic trees. Available from http://tree.bio.ed.ac.uk/software/figtree/, 2014.
101. Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst Biol. 2012;61(6):1061–7.
102. Gruber M, Söding J, Lupas AN. Comparative analysis of coiled-coil prediction methods. J Struct Biol. 2006;155(2):140–5.
103. Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol. 1999;292:195–202.
104. Buchan DWA, Jones DT. The PSIPRED protein analysis workbench: 20 years on. Nucleic Acids Res. 2019;47:W402–7.
105. Jones DT, Cozzetto D. DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinformatics. 2015;31:857–63.
106. Kosakovsky Pond SL, Frost SDW, Muse SV. HyPhy: hypothesis testing using phylogenies. Bioinformatics. 2005;21(5):676–9.
107. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. PLoS Genet. 2012;8(7):e1002764.
108. Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Mol Biol Evol. 2015;32(5):1342–53.
109. Labbé RM, Irimia M, Currie KW, Lin A, Zhu SJ, Brown DDR, et al. A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. Stem Cells. 2012;30:1734–45.
110. Tapial J, Ha KCH, Sterne-Weiler T, Gohr A, Braunschweig U, Hermoso-Pulido A, et al. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res. 2017;27:1759–68.
111. Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, et al. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 2016;44:W147–53.
112. Ferran JL, Ayad A, Merchán P, Morales-Delgado N, Sánchez-Arrones L, Alonso A, et al. Exploring brain genoarchitecture by single and double chromogenic in situ hybridization (ISH) and immunohistochemistry (IHC) on cryostat, paraffin, or floating sections. In: Hauptmann G, editor. In Situ Hybridization Methods, Neuromethods, vol. 99. New York Springer; 2015. p. 83–107.
113. Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK, et al. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. Nat Methods. 2015;12:982–8.
114. Casaca A, Nóvoa A, Mallo M. Hoxb6 can interfere with somitogenesis in the posterior embryo through a mechanism independent of its rib-promoting activity. Development. 2016;143:437–48.
115. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier L. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816–21.
116. Lobjois V, Bel-Vialar S, Trousse F, Pituello F. Forcing neural progenitor cells to cycle is insufficient to alter cell-fate decision and timing of neuronal differentiation in the spinal cord. Neural Dev. 2008;3:4.
117. Mirra S, Ulloa F, Gutierrez-Vallejo I, Martí E, Soriano E. Function of Armcx3 and Armc10/SVH genes in the regulation of progenitor proliferation and neural differentiation in the chicken spinal cord. Front Cell Neurosci. 2016;10:47.
118. Deacon RMJ. Assessing nest building in mice. Nat Protoc. 2006;1:1117–9.
119. Andrade-Talavera Y, Duque-Feria P, Paulsen O, Rodríguez-Moreno A. Presynaptic spike timing-dependent long-term depression in the mouse hippocampus. Cereb Cortex. 2016;26:3637–54.
120. González-Rodríguez P, Ugidos IF, Pérez-Rodríguez D, Anuncibay-Soto B, Santos-Galdiano M, Font-Belmonte L, et al. Brain-derived neurotrophic factor alleviates the oxidative stress induced by oxygen and glucose deprivation in an ex vivo brain slice model. J Cell Physiol. 2019;234:9592–604.
121. Hamburger V, Hamilton HL. A series of normal stages in the development of the chick embryo. Dev Dyn. 1992;195:231–72.
122. Megason SG, McMahon AP. A mitogen gradient of dorsal midline Wnts organizes growth in the CNS. Development. 2002;129:2087–98.

**Publisher's Note**

# Characterization of an eutherian gene cluster generated after transposon domestication identifies *Bex3* as relevant for advanced neurological functions

Enrique Navas-Pérez[†], Cristina Vicente-García[†], Serena Mirra[†], Demian Burguera, Noèlia Fernàndez-Castillo, José Luis Ferrán, Macarena López-Mayorga, Marta Alaiz-Noya, Irene Suárez-Pereira, Ester Antón-Galindo, Fausto Ulloa, Carlos Herrera-Úbeda, Pol Cuscó, Rafael Falcón-Moya, Antonio Rodríguez-Moreno, Salvatore D'Aniello, Bru Cormand, Gemma Marfany, Eduardo Soriano, Ángel M. Carrión, Jaime J. Carvajal[*], Jordi Garcia-Fernàndez[*]

[†] These authors contributed equally to this work.

[*] Corresponding authors. E-mails: jcarvajal@csic.es (J.J.C), jordigarcia@ub.edu (J.G.-F.).

## Supplementary information

**Table of contents:**

1

**Table S1** Coding genes putatively derived from transposable elements in the human and mouse genomes.

**Table S2** Altered expression of *BEX* and *TCEAL* genes in subjects with autism spectrum disorder or schizophrenia.

**Table S3** Enrichment of differential gene expression in *BEX* and *TCEAL* gene families in subjects with autism spectrum disorder or schizophrenia.

**Table S4** Primers and reconstructed gene sequences used in this work.

**Supplementary references**

2

**Fig. S1** The *Tceal7* gene is derived from the domestication of L1 retrotransposon fragments. Maximum-likelihood phylogenetic trees constructed using the nucleotidic sequences of ancient eutherian and metatherian L1 subfamilies, and the correspoding highlighted fragment of the *Tceal7* gene of three eutherian species: Hsa, *Homo sapiens*; Eca, *Equus caballus*; Laf, *Loxodonta africana*. The trees were inferred using IQ-TREE. The statistical support values are SH-aLRT and ultrafast bootstrap (UFBoot). The scale bar represents the expected number of nucleotide substitutions per site.
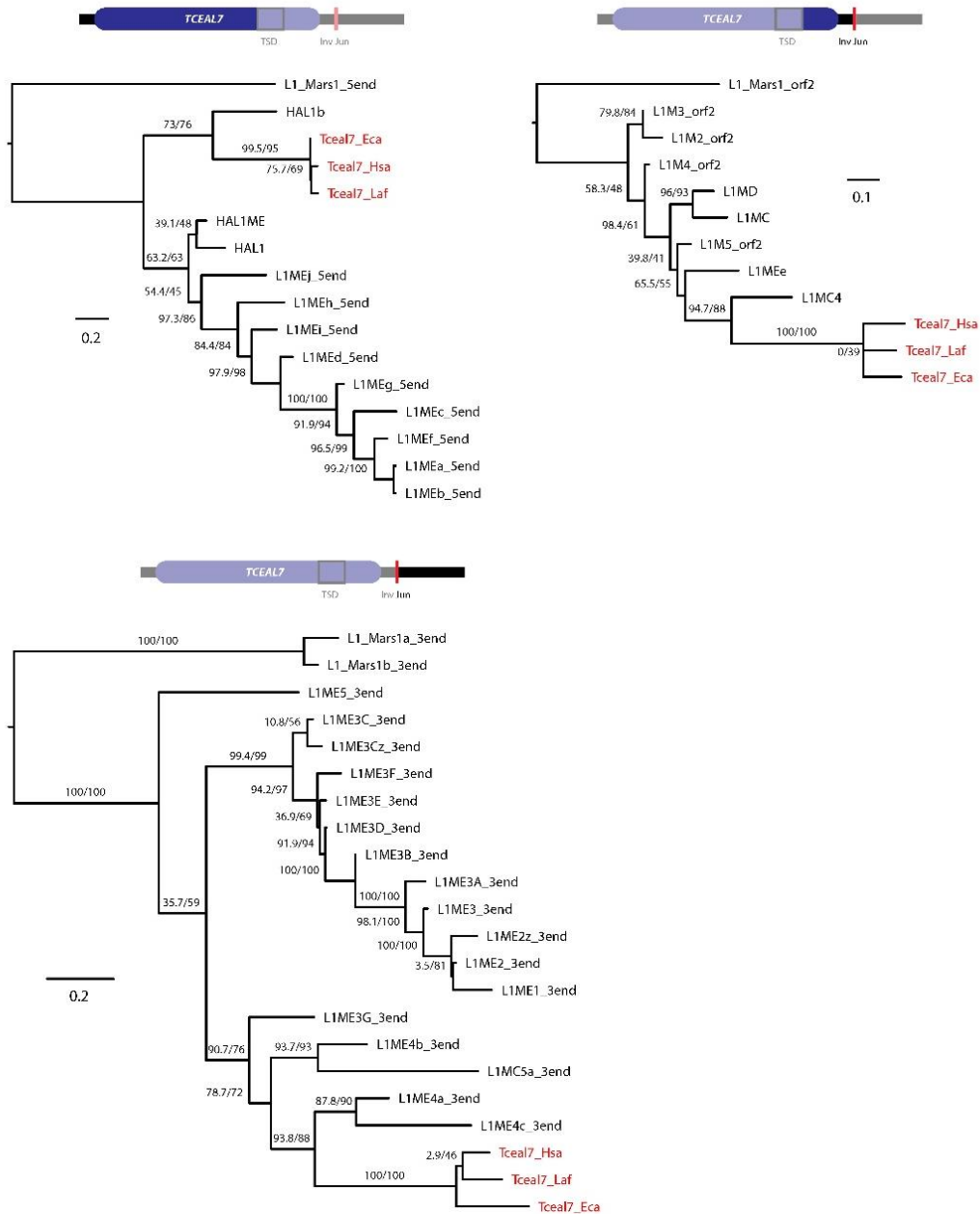


3

**Fig. S2** The *Bex/Tceal* gene cluster was established before the diversification of extant eutherians. Genes are represented with triangles and pseudogenes with empty triangles. The transparent pseudogenes in armadillo's row were found in the genome of another xenarthran: Hoffmann's two-toed sloth (*Choloepus hoffmanni*). The phylogenetic tree of eutherian species was constructed according to previous studies [1].
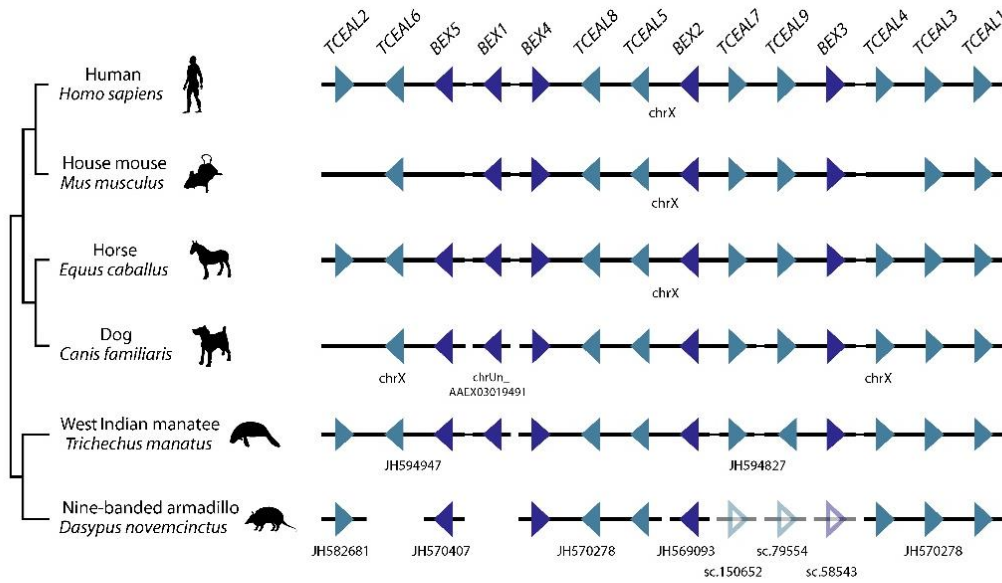
**Fig. S3** Highly diverged BEX/TCEAL proteins share a coiled-coil domain. **A** Protein alignment of human BEX and TCEAL proteins. Shaded in soft red and strong red are regions predicted by PCOILS to form coiled coils with a probability of 0.5-0.9 and > 0.9, respectively. **B** Maximum-likelihood phylogenetic tree constructed using the sequences of BEX and TCEAL proteins from eight eutherian species: dog, mouse, human, horse, cow, African bush elephant, nine-banded armadillo and West Indian manatee. The tree was inferred using IQ-TREE. The statistical support values are SH-aLRT and ultrafast bootstrap (UFBoot). The scale bar represents the expected number of substitutions per site.



5
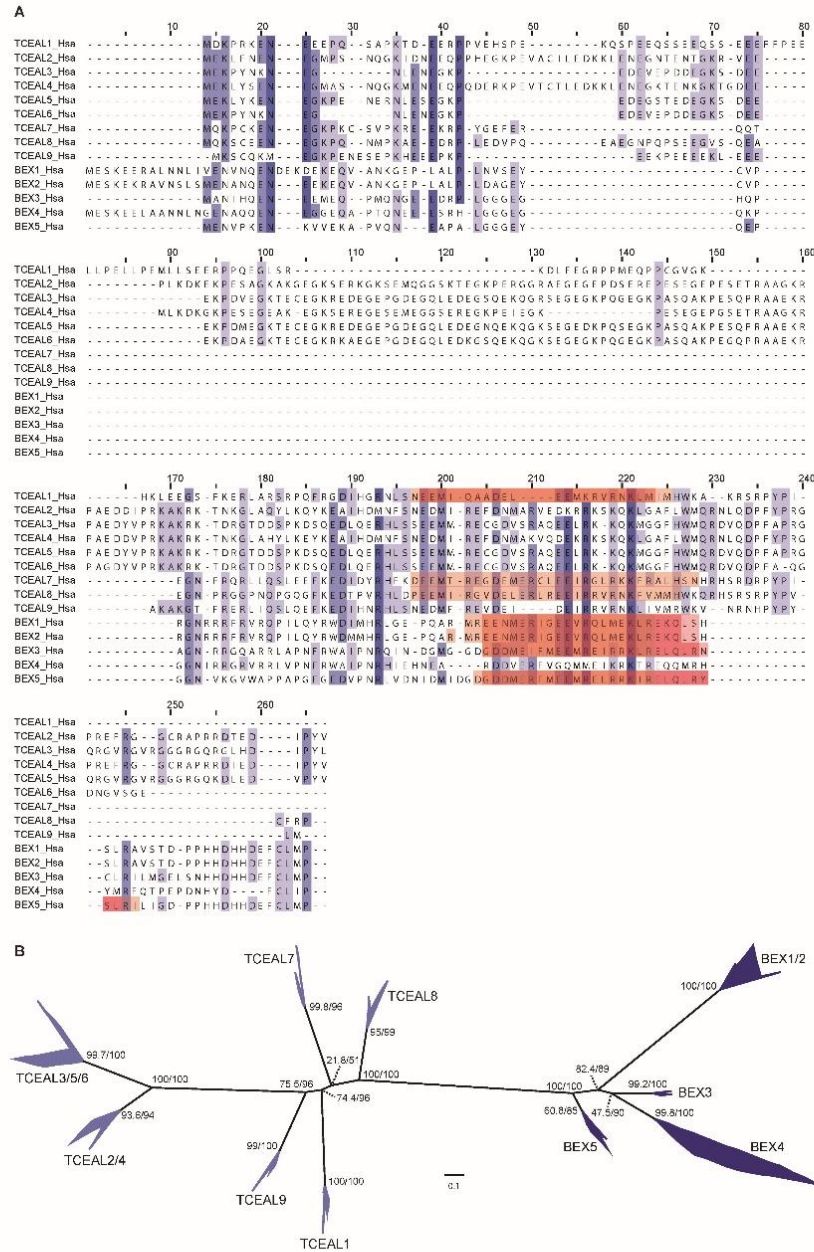
**Fig. S4** BEX/TCEAL proteins might have inherited some of their structural properties from the ancestral transposon. Protein disorder and α-helix predictions for the N-terminal end of HAL1b and for human BEX and TCEAL proteins.
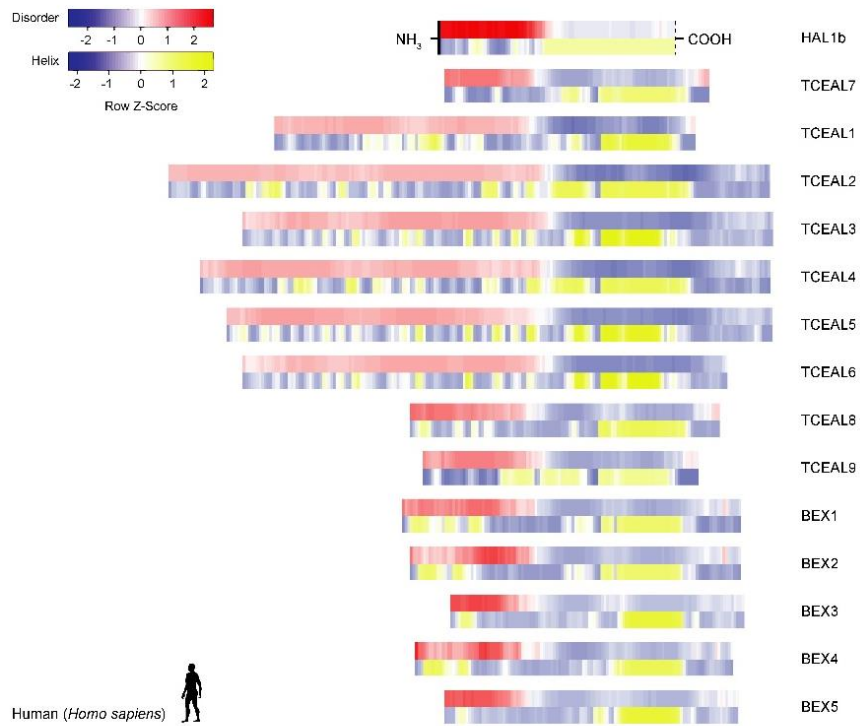
6

**Fig. S5** Selection pressure analyses reveal signatures of positive selection in the *Bex/Tceal* genes. MEME and aBSREL methods were used to infer selection pressures acting on **A** the *Bex* genes using *Tceal7* as an outgroup, and **B** the *Tceal* genes using *Bex5* as an outgroup. *Bex/Tceal* genes that homogenize their coding regions through gene conversion were excluded from the analysis. Amino acid sequences from eight eutherian species (Hsa, *Homo sapiens*; Eca, *Equus caballus*; Bta, *Bos taurus*; Laf, *Loxodonta africana*; Tma, *Trichechus manatus*; Cfa, *Canis familiaris*; Dno, *Dasypus novemcinctus*; Mmu, *Mus musculus*) were used to build the alignments and the phylogenetic trees. Branches under positive selection according to aBSREL ($P < 0.05$) are shown in red on the phylogenetic trees, and the ω1 and ω2 values indicate that there are two ω classes in a given branch with different selective pressures. Sites in the alignment under positive selection according to MEME ($P < 0.05$) are shown highlighted in grey on the right. Phylogenetic trees were inferred using IQ-TREE. The statistical support values are SH-aLRT and ultrafast bootstrap (UFBoot). The scale bar represents the expected number of substitutions per site.
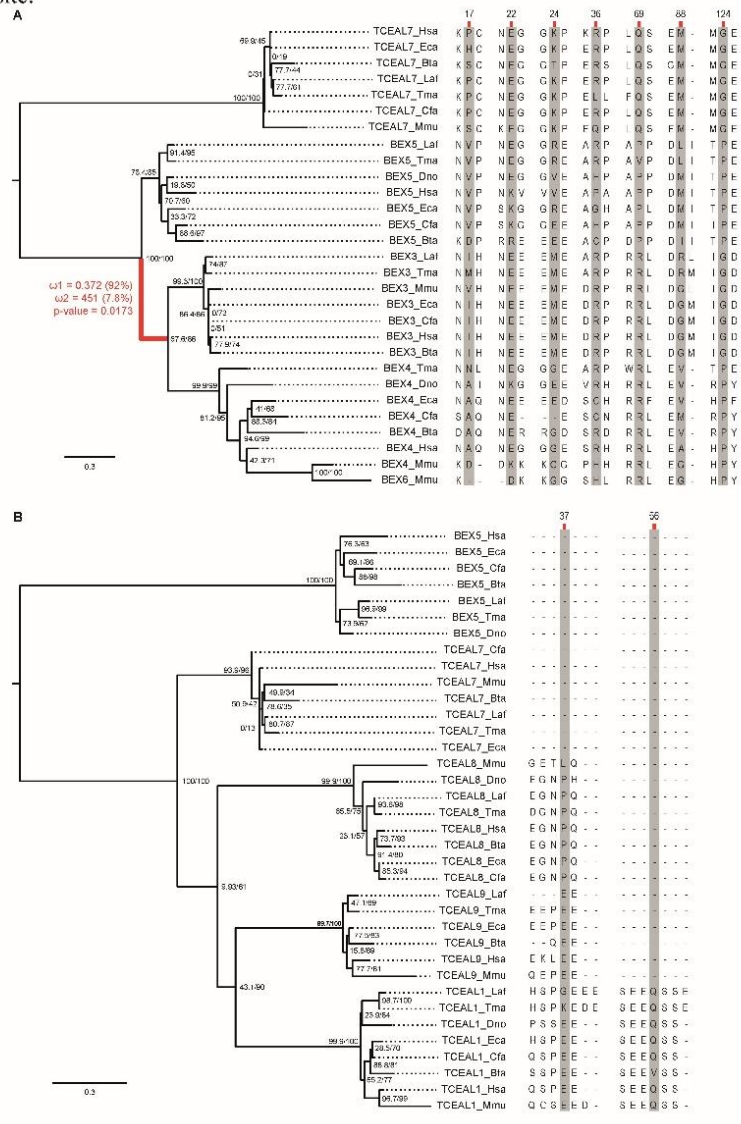
**Fig. S6** A BGW-like sequence was already present in the *GLA* promoter of the last therian common ancestor. Alignment of the *Hnrnph2* BGW motif and its orthologous region in marsupial species. Shaded in red, the conserved E-box sequence.



8
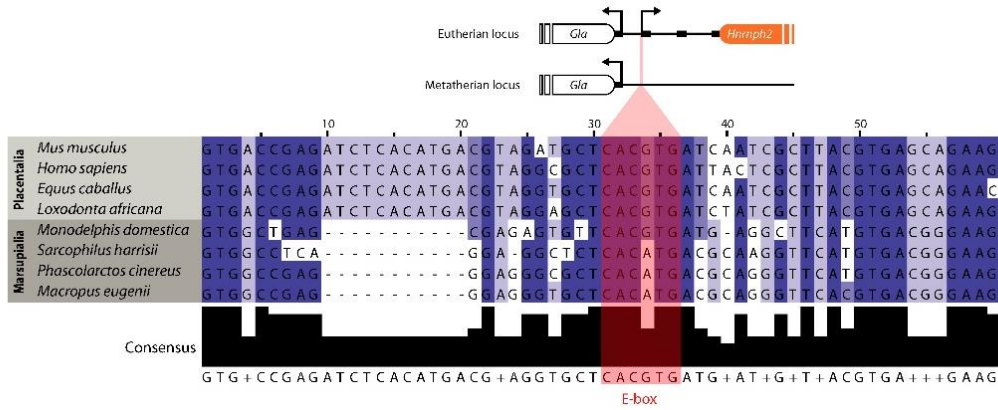
**Fig. S7** *Bex/Tceal* genes show tissue-enriched expression patterns during development. Expression patterns of *Bex/Tceal* genes using *in situ* hybridization (ISH) in E13.5 mouse embryos. Whole sagittal sections and details showing expression in specific tissues are shown. ca, cartilage; drg, dorsal root ganglion; ge, gut epithelium; gg gasserian ganglion; he, heart; ki, kidney; le, lens; li, liver; lu, lung; oe, olfactory epithelium; pe, pancreatic epithelium; re, retina; Rp, Rathke's pouch; se, stomach epithelium; sm, skeletal muscle; tg, thyroid gland; tr, thymic rudiment. Scale bar of whole sagittal sections: 1mm. Scale bar of details: 250μm.
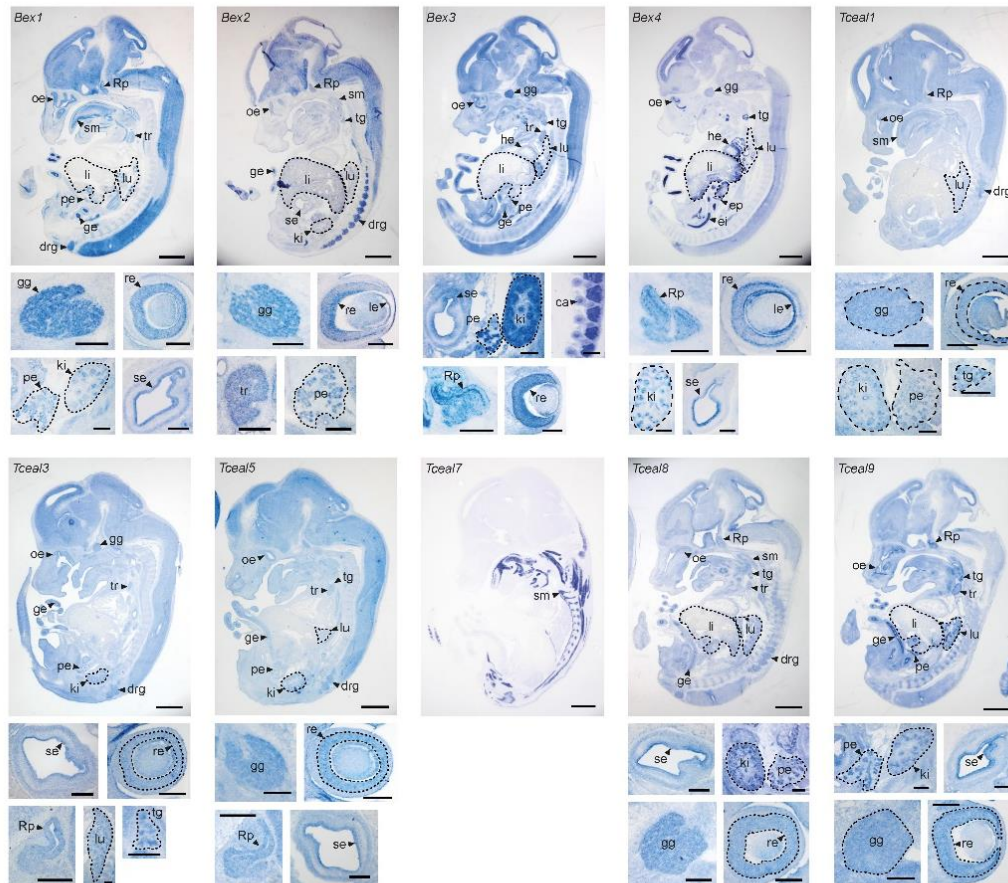


9

**Fig. S8** *Bex3* and *Tceal7* genes, but not the ancestral HALEX element, induce cell proliferation in chicken neural tube. **A-B** Representative transverse sections of neural tubes from embryos electroporated at HH12 with pCIG, pCIGBex3, pCIGTceal7 and pCIGHALEX vectors and analysed at 24 hpe by immunostaining. GFP (green) and Sox2 (red) stain, respectively, the electroporated cells and the neural progenitors. **C-D** The percentage of GFP-positive electroporated cells positive for BrdU or Sox2 increases in pCIGBex3 and pCIGTceal7 electroporated embryos. Cell counting was carried out on 10-17 pictures obtained from 4 to 7 different chick embryos per experimental condition. Data represent the mean ± SEM; * $P < 0.05$, ** $P < 0.01$, unpaired Student's t tests.
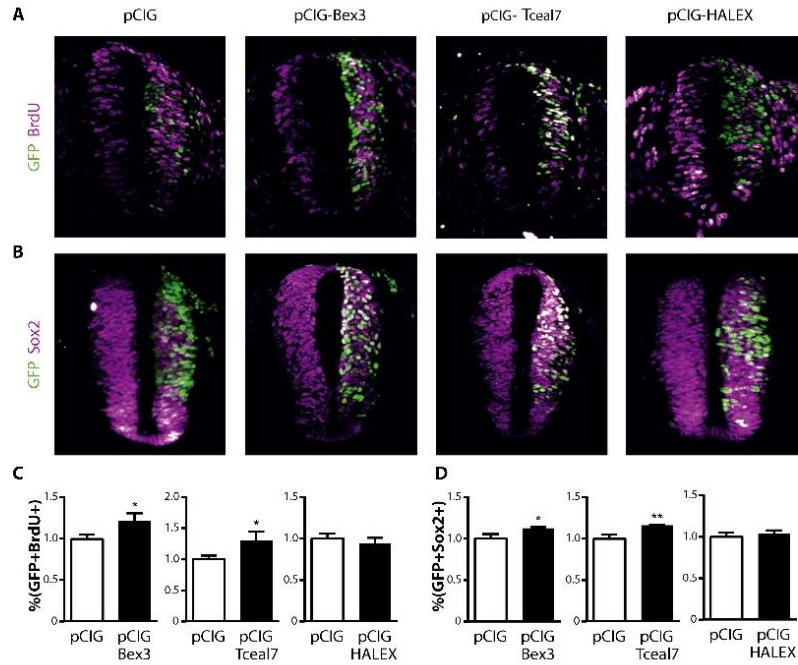
10

**Fig. S9** The deletions introduced using CRISPR-Cas9 technology can be observed in the mRNA expressed from the *Bex3* mutant alleles. Brain expression of *Bex3* and *actin* from three adult male wild-type and mutant animals from both lines.
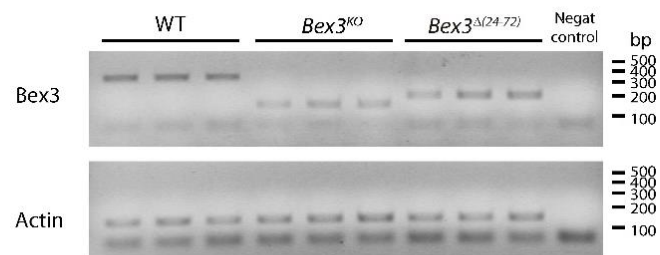
**Fig. S10** CRISPR-Cas9-generated *Bex3* mutant alleles show subtle skull abnormalities. Morphometric analyses of skulls from wild-type and mutant $Bex3^{KO}$ 6-week-old males **A** employing a total of 24 anatomical measurements **B** revealed that *Bex3* dysfunction led to cranial abnormalities in frontal bone and skull height **C**. Measurements were normalized to maximum skull length (D0) and expressed relative to controls (black horizontal line). Deviations of 5% with respect to controls are shown as dotted red and green horizontal lines. Results are presented as mean ± SEM ($N \geq 4$); * $P < 0.05$, one-way ANOVA.
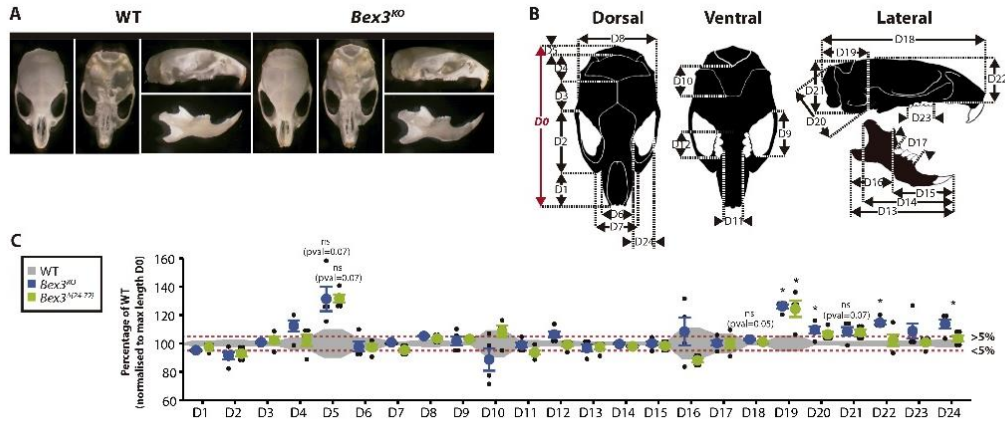
**Fig. S11** *Bex3* mutant mice show normal acoustic startle reflex. To analyze the acoustic startle reflex, latency of response, latency of peak and peak amplitude of the startle response were evaluated. Results are presented as mean ± SEM ($N$ = 6 per experimental group); one-way ANOVA.
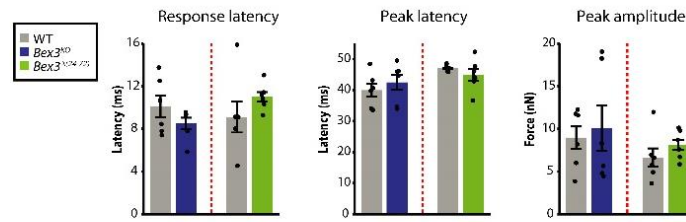
**Fig. S12** *Bex3* deficiency leads to aberrant mTOR signalling in the brain. **A** Western blot analyses of whole brain lysates of adult *Bex3*$^{\Delta(24-72)}$ mice revealed abnormal phosphorylation ratios of some mTORC1 targets. Representative images. **B** Quantification relative to wild-type. Results are presented as mean ± SEM ($N = 4$ per experimental group); * $P < 0.05$, non-parametric Mann–Whitney test.
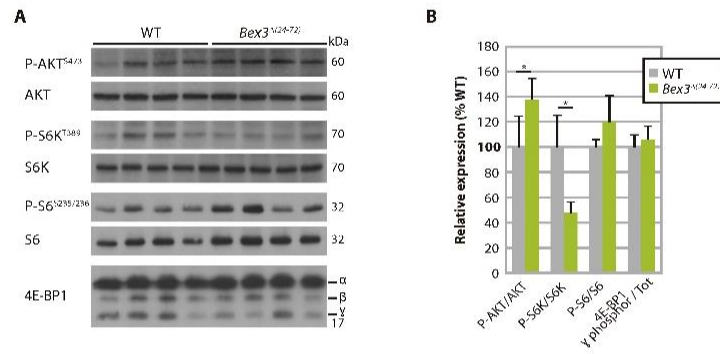
14

Table S1. Coding genes putatively derived from transposable elements in the human and mouse genomes.

| Gene ID | % overlap | KnownGene | txCdsPredict score | Domesticated transposon | Reported | Conservation |
|---|---|---|---|---|---|---|
| **Homo sapiens** | | | | | | |
| AK127846 | 100 | uc002qao.3 | 2124 | MER50-int (LTR retrotransposon) | - | Simiiformes |
| C5orf34 (ZBED8) | 100 | uc003lye.1 | 3097.83 | Charlie11 (Cut and Paste DNA transposon) | [2] | |
| ERV3-1 | 100 | uc011kdr.2 | 1829.5 | HERV3-int (LTR retrotransposon) | [3] | |
| ERVFRD-1 (Syncitin-2) | 57.05 | uc003mzt.3 | 2188 | MER50-int (LTR retrotransposon) | [4] | |
| ERVK6 | 99.93 | uc032zcp.1 | 1583.5 | HERVK-int (LTR retrotransposon) | [5] | |
| ERVMER34-1 | 53.81 | uc003gzs.3 | 1965.67 | MER34-int (LTR retrotransposon) | [6] | |
| ERVW-1 (Syncytin-1) | 100 | uc022ahe.2 | 1630.5 | HERV17-int (LTR retrotransposon) | [7] | |
| FAM200A | 94.02 | uc003urb.3 | 2929.5 | Charlie9 (Cut and Paste DNA transposon) | [8] | |
| FAM200B | 100 | uc003gof.4 | 2993 | Charlie9 (Cut and Paste DNA transposon) | [8] | |
| GTF2IRD2 | 55.99 | uc032zsc.1 | 4911 | Charlie8 (Cut and Paste DNA transposon) | [9] | |
| GTF2IRD2B | 100 | uc003ubt.3 | 4988.33 | Charlie8 (Cut and Paste DNA transposon) | [9] | |
| JRK | 90.57 | uc033cbj.1 | 3028.5 | Tigger10 (Cut and Paste DNA transposon) | [10] | |
| L1TD1 | 53.81 | uc001dae.5 | 5296.67 | L1MEd, HAL1ME (LINE retrotransposon) | [11] | |
| MGC12965 | 54.66 | uc001ovc.3 | 1148 | Tigger4 (Cut and Paste DNA transposon) | - | Primates |
| MOAP1 (PNMA4) | 96.5 | uc001ybj.3 | 2045.33 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA1 | 96.14 | uc001xor.1 | 2279 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA2 | 92.7 | uc003xez.2 | 2302 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA3 | 72.72 | uc033far.1 | 2872.33 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA3 | 73.94 | uc022cho.2 | 2854.33 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA5 | 60.76 | uc010ntw.3 | 2491.67 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA6A | 73.19 | uc022chq.2 | 2306.67 | MamGyp-int (LTR retrotransposon) | [12] | |
| PNMA6A | 73.19 | uc011myl.3 | 1972 | MamGyp-int (LTR retrotransposon) | [12] | |
| SETMAR | 71.84 | uc011asp.2 | 3181.83 | HSMAR1 (Cut and Paste DNA transposon) | [13] | |
| TCEAL7 | 76.32 | uc004ekc.2 | 804 | HAL1b (LINE retrotransposon) | | Eutheria |
| TIGD1 | 100 | uc002vsy.2 | 2366 | Tigger1 (LTR retrotransposon) | [14] | |
| ZBED5 | 97.94 | uc009ygh.3 | 3590.5 | Charlie14a (LTR retrotransposon) | [2] | |
| ZMYM5 | 51.11 | uc001umm.1 | 2329 | Zaphod2 (Cut and Paste DNA transposon) | [15] | |
| ZMYM6 (ZBED7) | 69.92 | uc031txb.1 | 3128.17 | Charlie10 (Cut and Paste DNA transposon) | [15] | |
| **Mus musculus** | | | | | | |
| BC033947 | 94.78 | uc011wnt.1 | 2345 | RLTR6-int (LTR retrotransposon) | - | Mus spretus |
| Bri1os | 79.18 | uc007ksg.1 | 947 | RLTR6-int (LTR retrotransposon) | - | Mus (genus) |
| Jrk | 92.72 | uc033gue.1 | 3457 | Tigger10 (Cut and Paste DNA transposon) | [10] | |
| Moap1 (Pnma4) | 96.13 | uc011yqp.2 | 1593.5 | MamGyp-int (LTR retrotransposon) | [12] | |
| Pnma1 | 96.14 | uc007ocl.1 | 2267 | MamGyp-int (LTR retrotransposon) | [12] | |
| Pnma2 | 92.45 | uc033grq.1 | 2386 | MamGyp-int (LTR retrotransposon) | [12] | |
| Pnma3 | 71.97 | uc009tla.1 | 2513 | MamGyp-int (LTR retrotransposon) | [12] | |
| Skint5 | 67.71 | uc009vcz.1 | 4936.5 | MurERV4-int (LTR retrotransposon) | - | Mus spretus |
| Zbed5 | 78.51 | uc008zto.1 | 2692.5 | Charlie14a (LTR retrotransposon) | [2] | |

15

Table S2. Altered expression of *BEX* and *TCEAL* genes in subjects with autism spectrum disorder or schizophrenia.

| Gene symbol | Disorder | Fold change | P value | FDR | Probe | Tissue | Study (PMID) or GEO ID | Sample, cases vs. controls |
|---|---|---|---|---|---|---|---|---|
| BEX1 | Autism | -1.60 | 3.59x10-4 | 2.03x10-2 | ILMN_1702637* | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| BEX1 | Autism | N/A | N/A | N/A | N/A | dorsolateral prefrontal cortex | [16] (27685936) | 34 ASD vs. 40 control |
| BEX1 | Autism | -1.21 | 2.26x10-2 | 1.20x10-1 | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| BEX1 | Autism | -1.43 | N/A | 4.19x10-3 | N/A | prefrontal and anterior cingulate cortex (L4 ) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.25 | N/A | 4.92x10-2 | N/A | prefrontal and anterior cingulate cortex (L5/6 ) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.74 | N/A | 1.47x10-3 | N/A | prefrontal and anterior cingulate cortex (Neu-NRGN-II) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.21 | N/A | 2.28x10-2 | N/A | prefrontal and anterior cingulate cortex (IN-PV) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.31 | N/A | 5.90x10-3 | N/A | prefrontal and anterior cingulate cortex (IN-SST) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.29 | N/A | 1.03x10-4 | N/A | prefrontal and anterior cingulate cortex (IN-VIP) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Autism | -1.24 | N/A | 6.95x10-3 | N/A | prefrontal and anterior cingulate cortex (IN-SV2C) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX1 | Schizophrenia | -1.08 | 8.22x10-4 | 5.90x10-2 | 8171201 | cerebellum | GSE35978 | 41 scz vs. 50 control |
| BEX1 | Schizophrenia | -1.33 | 1.42x10-4 | 1.18x10-2 | 218332_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| BEX1 | Schizophrenia | -1.06 | 3.38x10-2 | 3.34x10-1 | 218332_at | prefrontal cortex | GSE53987 | 15 scz vs. 19 control |
| BEX1 | Schizophrenia | -1.14 | 2.84x10-2 | 2.92x10-1 | 218332_at | striatum | GSE53987 | 18 scz vs. 18 control |
| BEX1 | Schizophrenia | -1.07 | 8.57x10-4 | 5.92x10-2 | 11719475_a_at | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| BEX1 | Schizophrenia | -1.09 | N/A | 4.65x10-3 | 218332_at | prefrontal cortex | [19] (26818902) | 122 scz vs. 124 control |
| BEX1 | Schizophrenia | N/A | 4.32x10-4 | 9.26x10-2 | N/A | anterior cingulate cortex | [20] (25113377) | 31 scz vs. 26 control |
| BEX2 | Autism | -1.53 | 5.21x10-4 | 2.31x10-2 | ILMN_2181892 | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| BEX2 | Autism | -1.20 | 4.12x10-2 | 1.67x10-1 | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| BEX2 | Schizophrenia | -1.26 | 1.21x10-3 | 2.59x10-2 | 224367_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| BEX2 | Schizophrenia | -1.09 | 1.15x10-2 | 1.97x10-1 | 11718848_a_at | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| BEX3 | Autism | -1.48 | 1.43x10-2 | 1.25x10-1 | ILMN_1729208* | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| BEX3 | Autism | -1.18 | 3.66x10-3 | 5.15x10-2 | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| BEX3 | Autism | -1.05 | 2.35x10-2 | 1.61x10-1 | N/A | frontal and temporal cortex | [21] (30545856) | 51 ASD vs. 936 control |
| BEX3 | Autism | -1.25 | N/A | 4.50x10-2 | N/A | prefrontal and anterior cingulate cortex (L4 ) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX3 | Autism | -1.62 | N/A | 4.12x10-2 | N/A | prefrontal and anterior cingulate cortex (Neu-NRGN-II) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX3 | Autism | -1.29 | N/A | 5.45x10-4 | N/A | prefrontal and anterior cingulate cortex (IN-VIP) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX3 | Autism | -1.34 | N/A | 3.81x10-3 | N/A | prefrontal and anterior cingulate cortex (IN-SV2C) | [18] (31097668 ) | 15 ASD vs. 16 control |
| BEX3 | Schizophrenia | -1.03 | 4.77x10-2 | 3.39x10-1 | 8169028 | cerebellum | GSE35978 | 44 scz vs. 50 control |
| BEX3 | Schizophrenia | -1.28 | 1.88x10-4 | 1.28x10-2 | 217963_s_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| BEX3 | Schizophrenia | -1.06 | 4.56x10-2 | 3.76x10-1 | 217963_s_at | prefrontal cortex | GSE53987 | 15 scz vs. 19 control |
| BEX3 | Schizophrenia | -1.11 | 1.81x10-2 | 2.67x10-1 | 217963_s_at | striatum | GSE53987 | 18 scz vs. 18 control |
| BEX3 | Schizophrenia | -1.28 | 5.00x10-3 | N/A | 217963_s_at | dorsolateral prefrontal cortex | [22] (24886351) | 8 scz vs. 7 control |
| BEX3 | Schizophrenia | -1.06 | N/A | 5.41x10-3 | 217963_s_at | prefrontal cortex | [19] (26818902) | 122 scz vs. 124 control |
| BEX4 | Autism | -1.65 | 6.17x10-3 | 7.82x10-2 | ILMN_2351638 | cerebellum | GSE38322 | 14 ASD vs. 12 control |

16

| Gene symbol | Disorder | Fold change | P value | FDR | Probe | Tissue | Study (PMID) or GEO ID | Sample, cases vs. controls |
|---|---|---|---|---|---|---|---|---|
| BEX4 | Autism | 1.10 | 4.26x10-2 | 6.72x10-1 | ILMN_1773504 | occipital cortex | GSE38322 | 6 ASD vs. 4 control |
| BEX4 | Autism | -1.14 | 3.45x10-2 | 4.68x10-1 | N/A | cerebellum | [17] (27919067) | 48 ASD vs. 49 control |
| BEX4 | Autism | -1.11 | 2.43x10-2 | 1.25x10-1 | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| BEX4 | Autism | -1.11 | 3.00x10-2 | N/A | ILMN_1804798 | prefrontal and temporal cortex | [23] (29859039) | 15 ASD vs. 16 control |
| BEX4 | Schizophrenia | -1.30 | 4.70x10-4 | **1.75x10-2** | 215440_s_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| BEX4 | Schizophrenia | -1.16 | 2.46x10-2 | 2.83x10-1 | 215440_s_at | striatum | GSE53987 | 18 scz vs. 18 control |
| BEX5 | Autism | -1.93 | 3.78x10-4 | **2.80x10-2** | ILMN_1806473 | cerebellum and occipital cortex | [24] (22984548) | 9 ASD vs. 9 control |
| BEX5 | Autism | -2.07 | 8.63x10-5 | **1.32x10-2** | ILMN_1806473 | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| BEX5 | Autism | -2.30 | 1.36x10-2 | 5.78x10-1 | ILMN_1806473 | occipital cortex | GSE38322 | 6 ASD vs. 4 control |
| BEX5 | Autism | -1.27 | 2.09x10-2 | 4.44x10-1 | N/A | cerebellum | [17] (27919067) | 48 ASD vs. 49 control |
| BEX5 | Autism | -1.29 | 3.35x10-3 | **4.92x10-2** | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| BEX5 | Autism | -1.49 | 3.53x10-2 | N/A | ILMN_1806473 | prefrontal and temporal cortex | [23] (29859039) | 15 ASD vs. 16 control |
| BEX5 | Autism | -1.10 | 2.70x10-2 | 1.74x10-1 | N/A | frontal and temporal cortex | [21] (30545856) | 51 ASD vs. 936 control |
| BEX5 | Schizophrenia | -1.13 | 1.23x10-4 | **2.58x10-2** | 8174141 | cerebellum | GSE35978 | 44 scz vs. 50 control |
| BEX5 | Schizophrenia | -1.40 | 7.24x10-4 | **2.05x10-2** | 229963_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| BEX5 | Schizophrenia | -1.10 | 5.00x10-2 | 3.86x10-1 | 229963_at | prefrontal cortex | GSE53987 | 15 scz vs. 19 control |
| BEX5 | Schizophrenia | -1.24 | 1.10x10-2 | 2.50x10-1 | 229963_at | striatum | GSE53987 | 18 scz vs. 18 control |
| BEX5 | Schizophrenia | -1.15 | 8.11x10-4 | **5.71x10-2** | 11725493_s_at | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| TCEAL1 | Autism | -1.62 | 1.86x10-4 | **1.65x10-2** | ILMN_2398403* | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| TCEAL1 | Autism | -1.65 | 3.94x10-2 | 6.62x10-1 | ILMN_2398408* | occipital cortex | GSE38322 | 6 ASD vs. 4 control |
| TCEAL1 | Autism | -1.16 | 1.55x10-2 | 4.37x10-1 | N/A | cerebellum | [17] (27919067) | 48 ASD vs. 49 control |
| TCEAL1 | Autism | -1.10 | 2.91x10-2 | 1.38x10-1 | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| TCEAL1 | Schizophrenia | -1.06 | 2.90x10-3 | 1.06x10-1 | 8169049 | cerebellum | GSE35978 | 44 scz vs. 50 control |
| TCEAL1 | Schizophrenia | -1.17 | 1.08x10-2 | **8.21x10-2** | 204045_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL1 | Schizophrenia | -1.18 | 4.48x10-3 | 2.30x10-1 | 204045_at | striatum | GSE53987 | 18 scz vs. 18 control |
| TCEAL2 | Autism | -1.12 | N/A | **1.52x10-3** | N/A | prefrontal and anterior cingulate cortex (L4) | [18] (31097668) | 15 ASD vs. 16 control |
| TCEAL2 | Schizophrenia | -1.13 | 4.10x10-2 | 1.77x10-1 | 211276_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL2 | Schizophrenia | -1.07 | 2.60x10-2 | 2.75x10-1 | 11727253_at | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| TCEAL2 | Schizophrenia | 1.03 | 2.05x10-3 | **1.61x10-2** | N/A | frontal and temporal cortex | [21] (30545856) | 559 scz vs. 936 control |
| TCEAL3 | Autism | N/A | N/A | N/A | N/A | dorsolateral prefrontal cortex | [16] (27685936) | 34 ASD vs. 40 control |
| TCEAL3 | Schizophrenia | -1.15 | 2.34x10-3 | **3.59x10-2** | 227279_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL3 | Schizophrenia | -1.25 | 4.52x10-5 | **1.16x10-2** | 11725131_x_at* | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| TCEAL4 | Autism | -1.10 | 4.02x10-2 | N/A | ILMN_1748625 | prefrontal and temporal cortex | [23] (29859039) | 15 ASD vs. 16 control |
| TCEAL4 | Schizophrenia | -1.17 | 8.89x10-4 | **2.23x10-2** | 202371_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL4 | Schizophrenia | -1.07 | 1.76x10-2 | 2.65x10-1 | 202371_at | striatum | GSE53987 | 18 scz vs. 18 control |
| TCEAL4 | Schizophrenia | -1.13 | 4.27x10-3 | 1.28x10-1 | 11745071_a_at | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| TCEAL5 | Autism | -1.17 | 3.10x10-2 | 1.91x10-1 | ILMN_1749073 | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| TCEAL6 | Autism | -1.48 | 1.82x10-2 | 1.42x10-1 | ILMN_1729165* | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| TCEAL6 | Schizophrenia | -1.47 | 6.37x10-6 | **4.25x10-3** | 11737036_at* | dorsolateral prefrontal cortex | GSE87610 | 65 scz vs. 72 control |
| TCEAL6 | Schizophrenia | 1.03 | 1.68x10-2 | **7.69x10-2** | N/A | frontal and temporal cortex | [21] (30545856) | 559 scz vs. 936 control |
| TCEAL7 | Autism | -1.71 | 1.30x10-3 | **3.41x10-2** | ILMN_1753525 | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| TCEAL7 | Autism | -1.24 | 1.69x10-2 | 5.86x10-1 | ILMN_2084043 | occipital cortex | GSE38322 | 6 ASD vs. 4 control |
| TCEAL7 | Autism | -1.18 | 3.06x10-3 | **4.76x10-2** | N/A | frontal and temporal cortex | [17] (27919067) | 48 ASD vs. 49 control |
| TCEAL7 | Autism | -1.06 | 4.42x10-2 | 2.29x10-1 | N/A | frontal and temporal cortex | [21] (30545856) | 51 ASD vs. 936 control |
| TCEAL7 | Schizophrenia | -1.09 | 4.20x10-3 | 1.23x10-1 | 8169015 | cerebellum | GSE35978 | 44 scz vs. 50 control |
| TCEAL7 | Schizophrenia | -1.35 | 1.23x10-4 | **1.12x10-2** | 227705_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL7 | Schizophrenia | -1.14 | 2.46x10-2 | 2.83x10-1 | 227705_at | striatum | GSE53987 | 18 scz vs. 18 control |

| Gene symbol | Disorder | Fold change | P value | FDR | Probe | Tissue | Study (PMID) or GEO ID | Sample, cases vs. controls |
|---|---|---|---|---|---|---|---|---|
| TCEAL7 | Schizophrenia | -1.04 | 2.36x10-5 | **5.21x10-4** | N/A | frontal and temporal cortex | [21] (30545856) | 559 scz vs. 936 control |
| TCEAL8 | Autism | -1.36 | 1.05x10-3 | **3.04x10-2** | ILMN_2402272* | cerebellum | GSE38322 | 14 ASD vs. 12 control |
| TCEAL8 | Autism | -1.56 | 5.82x10-3 | 5.53x10-1 | ILMN_2402272* | occipital cortex | GSE38322 | 6 ASD vs. 4 control |
| TCEAL8 | Autism | -1.13 | 2.36x10-2 | 4.49x10-1 | N/A | cerebellum | [17] (27919067) | 48 ASD vs. 49 control |
| TCEAL8 | Autism | -1.11 | 2.74x10-2 | N/A | ILMN_2402272 | prefrontal and temporal cortex | [23] (29859039) | 15 ASD vs. 16 control |
| TCEAL8 | Schizophrenia | -1.24 | 1.48x10-3 | **2.86x10-2** | 224819_at | hippocampus | GSE53987 | 15 scz vs. 18 control |
| TCEAL8 | Schizophrenia | -1.18 | 1.69x10-2 | 2.63x10-1 | 224819_at | striatum | GSE53987 | 18 scz vs. 18 control |
| TCEAL9 | Schizophrenia | 1.15 | 3.12x10-3 | 1.59x10-1 | 217975_at | prefrontal cortex | GSE53987 | 15 scz vs. 19 control |

N/A, data not available; FDR, False Discovery Rate; scz, schizophrenia cases; ASD, autism spectrum disorder cases; L4, layer 4 excitatory neurons; L5/6, layer 5/6 corticofugal projection neurons; Neu-NRGN, neurogranin–expressing neurons; IN-PV, parvalbumin interneurons; IN-SST, somatostatin interneurons;IN-VIP, vasoactive intestinal polypeptide–expressing interneurons; IN-SV2C, synaptic vesicle glycoprotein 2C-expressing interneurons. Experiments overcoming multiple testing corrections at 10% FDR are highlighted in bold. *Genes showing significant differential expression in several independent probe sets, data shown corresponding to the probe with the highest FC.

**Table S3. Enrichment of differential gene expression in *BEX* and *TCEAL* gene families in subjects with autism spectrum disorder or schizophrenia.**

| Study (PMID) or GEO ID | Disorder | Tissue | # genes analyzed | # DE genes | BEX genes | TCEAL genes | TOTAL | BEX enrichment (P value) | TCEAL enrichment (P value) | BEX+TCEAL enrichment (P value) |
|---|---|---|---|---|---|---|---|---|---|---|
| [21] (30545856) | Autism | frontal and temporal cortex | 25772 | 5250 | 2 | 1 | 1 | **1.10x10-2** | 1 | **3.04x10-2** |
| [24] (22984548) | Autism | cerebellum and occipital cortex | 18626 | 41 | 1 | 0 | 7 | **2.17x10-3** | 7.92x10-1 | 8.39x10-2 |
| GSE38322 | Autism | cerebellum | 20762 | 6312 | 5 | 5 | 4 | 8.34x10-2 | 2.30x10-1 | **4.60x10-2** |
| GSE38322 | Autism | occipital cortex | 20762 | 2341 | 2 | 3 | 4 | 2.19x10-1 | 4.96x10-1 | 2.30x10-1 |
| [16] (27685936) | Autism | prefrontal cortex | 12557 | 1775 | 1 | 1 | 10 | **2.59x10-3** | 1.04x10-1 | **1.86x10-3** |
| [17] (27919067) | Autism | cerebellum | 13190 | 1337 | 2 | 2 | 5 | 1.01x10-1 | 7.14x10-2 | **1.52x10-2** |
| [17] (27919067) | Autism | frontal and temporal cortex | 13543 | 3973 | 5 | 2 | 2 | 5.33x10-1 | 7.46x10-1 | 6.09x10-1 |
| [23] (29859039) | Autism | frontal and temporal cortex | 12632 | 2254 | 2 | 2 | 3 | **3.99x10-2** | 4.82x10-1 | 8.06x10-2 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (L4 ) | 1391 | 94 | 2 | 1 | 3 | **3.95x10-2** | 4.68x10-1 | 6.31x10-2 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (L5/6 ) | 1391 | 8 | 1 | 0 | 1 | **2.85x10-2** | 1 | 7.79x10-2 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (Neu-NRGN-II) | 1391 | 25 | 2 | 0 | 2 | **3.00x10-3** | 1 | **2.47x10-2** |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (IN-PV) | 1391 | 27 | 1 | 0 | 1 | 9.35x10-2 | 1 | 2.41x10-1 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (IN-SST) | 1391 | 27 | 1 | 0 | 1 | 9.35x10-2 | 1 | 2.41x10-1 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (IN-VIP) | 1391 | 78 | 2 | 0 | 2 | **2.78x10-2** | 1 | 1.83x10-1 |
| [18] (31097668) | Autism | prefrontal and anterior cingulate cortex (IN-SV2C) | 1391 | 32 | 2 | 0 | 2 | **4.91x10-3** | 1 | **3.93x10-2** |
| [21] (30545856) | Schizophrenia | frontal and temporal cortex | 25772 | 7840 | 0 | 3 | 5 | **9.70x10-5** | **8.32x10-4** | **3.36x10-7** |
| GSE35978 | Schizophrenia | cerebellum | 20359 | 3545 | 3 | 2 | 11 | **1.43x10-2** | 1.55x10-1 | **8.84x10-3** |
| GSE53987 | Schizophrenia | hippocampus | 22187 | 3497 | 5 | 6 | 8 | 8.27x10-2 | 9.04x10-1 | 4.04x10-1 |
| GSE53987 | Schizophrenia | prefrontal cortex | 22187 | 5083 | 3 | 1 | 4 | 5.06x10-1 | 1 | 8.61x10-1 |
| GSE53987 | Schizophrenia | striatum | 22187 | 5416 | 4 | 4 | 1 | **9.00x10-3** | 1 | 6.82x10-2 |
| GSE87610 | Schizophrenia | dorsolateral prefrontal cortex | 20034 | 5258 | 3 | 4 | 2 | **4.74x10-2** | 1 | 1.27x10-1 |
| [22] (24886351) | Schizophrenia | dorsolateral prefrontal cortex | 20955 | 2757 probes | 1 | 0 | 1 | **2.22x10-2** | **2.71x10-2** | **1.48x10-3** |
| [19] (26818902) | Schizophrenia | prefrontal cortex | 22277 probes | 690 | 2 | 0 | 7 | 2.70x10-1 | 8.71x10-1 | 5.66x10-1 |
| [20] (25113377) | Schizophrenia | anterior cingulate cortex | 14454 | 105 | 1 | 0 | 3 | 1 | 5.48x10-1 | 8.47x10-1 |

DE genes: differentially expressed genes. Enrichment calculated using an hypergeometric test for the 5 BEX genes and the 9 TCEAL genes. Significant *P* values are highlighted in bold.

**Table S4. Primers and reconstructed gene sequences used in this work.**

| Experiment | Name | Sequence |
|---|---|---|
| Generation of CRISPR-Cas9 sgRNA guides targeting *Bex3* | sgRNA1-F | agggGGACAGGAAGACCGCCCTGT |
| | sgRNA1-R | aaacACAGGGCGGTCTTCCTGTCC |
| | sgRNA2-F | agggCAGATGAATGACGGGTT |
| | sgRNA2-R | aaacAACCCGTCATTCATCTG |
| Template preparation for IDA // Genotyping of *Bex3* mutant lines | Bex3F-Ext | CCTGTCTAGGACCCCTGTGA |
| | Bex3R-Ext | GCGGGAGTCACAGTATGGAT |
| | Bex3F-Int | AGCCCCCACTCCACACTACT |
| | Bex3R-Int | TGGTGATCGTGGTGGTTAGA |
| RT-PCR *Bex3* | Bex3_RT_F | CCAATGTCCACCAGGAAAAC |
| | Bex3_RT_R | AGGCATAAGGCAGAATTCATCA |
| RT-PCR *actin* | act_RT_F | GGCTGTATTCCCCTCCATCG |
| | act_RT_R | CCAGTTGGTAACAATGCCATG |
| Riboprobe for *in situ* hyridization of *Bex1* | Bex1F | CATCATGACCACCATGATGAG |
| | Bex1R | GGTTCACAATAGGTAATACGG |
| Riboprobe for *in situ* hyridization of *Bex2* | Bex2F | CATCATGACCACCATGATGAG |
| | Bex2R | GGTTCACAATATATACTGAGC |
| Riboprobe for *in situ* hyridization of *Bex3* | Bex3F | CGAAGAGATGGAGCAGCCCC |
| | Bex3R | CATGCTAATGGGCAACACTG |
| Riboprobe for *in situ* hyridization of *Bex4* | Bex4F | GGCAAGGATAGGCCCAGGAG |
| | Bex4R | ATGATTGTCAGGTTCCGGGG |
| Riboprobe for *in situ* hyridization of *Tceal1* | Tceal1F | CAGCCTGCAGTGGAGCAGTC |
| | Tceal1R | TAAGGGCGGCTCCGTTTTGC |
| Riboprobe for *in situ* hyridization of *Tceal3* | Tceal3F | GTTCACGTGACCGGCCAGGC |
| | Tceal3R | GACAGGCCTTTGCCTTCTCC |
| Riboprobe for *in situ* hyridization of *Tceal5* | Tceal5F | GAAGTCTCTCTCTTCCAGGT |
| | Tceal5R | CCATGGTGCAATTAGTCTTTG |
| Riboprobe for *in situ* hyridization of *Tceal7* | Tceal7F | AGCCCAAGGGCAGTGAGGCA |
| | Tceal7R | CAGGTGACAGTTGACGTGCC |
| Riboprobe for *in situ* hyridization of *Tceal8* | Tceal8F | CGAGAACGAAGGAACACCCC |
| | Tceal8R | AAGCCAGAGGGAATTCCAGG |
| Riboprobe for *in situ* hyridization of *Tceal9* | Tceal9F | GGAGGATGAGCCAAAGCCTG |
| | Tceal9R | CATACAAAGGTACTCCTGTC |
| HALEX ancestral gene reconstruction | HALEX | ATGCAGACCTATTATTCAAAAGAGGAGCGGGGGAAAGAAGAAAGACGGCACTCAATGC AGAAGACCGACGAGGAACCAAGCCTGGAAGAGAATATCACCCAGGGCACAGAGGAAA ACAGCTCCCAGGTGCTGCACGCCATCGAAAAGCTGAACAAAAACATGAACTCTATCAAG CAGGAGCTGAAAGACGAAATGATTAAGCAGCAGAATGAGATGAAAAGAGAGATCGCTG AACTGAGGAAGCAGATTGAAAACCAGAACAACATCCTGCACTCCAATCACCGACATTCC CGAGACCTGCTGTACCTG |

**Supplementary references**

1. Tarver JE, Dos Reis M, Mirarab S, Moran RJ, Parker S, O'Reilly JE, et al. The interrelationships of placental mammals and the limits of phylogenetic inference. Genome Biol Evol. 2016;8:330–44.

2. Aravind L. The BED finger, a novel DNA-binding domain in chromatin-boundary-element-binding proteins and transposases. Trends Biochem Sci. 2000;25:421–3.

3. Boyd MT, Bax CMR, Bax BE, Bloxam DL, Weiss RA. The Human Endogenous Retrovirus ERV-3 Is Upregulated in Differentiating Placental Trophoblast Cells. Virology. 1993;196:905–9.

4. Blaise S, de Parseval N, Bénit L, Heidmann T. Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. Proc Natl Acad Sci U S A. 2003;100:13013–8.

5. Mayer J, Sauter M, Racz A, Scherer D, Mueller-Lantzsch N, Meese E. An almost-intact human endogenous retrovirus K on human chromosome 7. Nat Genet. 1999;21:257–8.

6. Heidmann O, Béguin A, Paternina J, Berthier R, Deloger M, Bawa O, et al. HEMO, an ancestral endogenous retroviral envelope protein shed in the blood of pregnant women and expressed in pluripotent stem cells and tumors. Proc Natl Acad Sci U S A. 2017;114:E6642–51.

7. Mi S, Lee X, Li X, Veldman GM, Finnerty H, Racie L, et al. McCoy, Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature. 2000;403:785–9.

8. Smit AF. Interspersed repeats and other mementos of transposable elements in mammalian genomes. Curr Opin Genet Dev. 1999;9:657–63.

9. Tipney HJ, Hinsley TA, Brass A, Metcalfe K, Donai D, Tassabehji M. Isolation and characterisation of GTF21RD2, a novel fusion gene and member of the TFII-I family of transcription factors, deleted in Williams-Beuren syndrome. Eur J Hum Genet. 2004;12:551–60.

10. Toth M, Grimsby J, Buzsaki G, Donovan GP. Epileptic seizures caused by inactivation of a novel gene, jerky, related to centromere binding protein-B in transgenic mice. Nat Genet. 1995;11:71–5.

21

11. McLaughlin RN, Young JM, Yang L, Neme R, Wichman HA, Malik HS. Positive selection and multiple losses of the LINE-1-derived L1TD1 gene in mammals suggest a dual role in genome defense and pluripotency. PLoS Genet. 2014;10:e1004531.

12. Campillos M, Doerks T, Shah PK, Bork P. Computational characterization of multiple Gag-like human proteins. Trends Genet. 2006;22:585–9.

13. Cordaux R, Udit S, Batzer MA, Feschotte C. Birth of a chimeric primate gene by capture of the transposase gene from a mobile element. Proc Natl Acad Sci U S A. 2006;103:8101–6.

14. Smit AFA, Riggs AD. Tiggers and other DNA transposon fossils in the human genome. Proc Natl Acad Sci U S A. 1996;93:1443–8.

15. Kojima KK, Jurka J. Crypton transposons: Identification of new diverse families and ancient domestication events. Mob DNA. 2011;2:12.

16. Liu X, Han D, Somel M, Jiang X, Hu H, Guijarro P, et al. Disruption of an Evolutionarily Novel Synaptic Expression Pattern in Autism. PLoS Biol. 2016;14:e1002558.

17. Parikshak NN, Swarup V, Belgard TG, Irimia M, Ramaswami G, Gandal MJ, et al. Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. Nature. 2016;540:423–7.

18. Velmeshev D, Schirmer L, Jung D, Haeussler M, Perez Y, Mayer S, et al. Single-cell genomics identifies cell type–specific molecular changes in autism. Science. 2019;364:685–9.

19. Qin W, Liu C, Sodhi M, Lu H. Meta-analysis of sex differences in gene expression in schizophrenia. BMC Syst Biol. 2016;10:9.

20. Zhao Z, Xu J, Chen J, Kim S, Reimers M, Bacanu S-A, et al. Transcriptome sequencing and genome-wide association analyses reveal lysosomal function and actin cytoskeleton remodeling in schizophrenia and bipolar disorder. Mol Psychiatry. 2015;20:563–72.

22

21. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. Science. 2018;362:pii:eaat8127.

22. Hagihara H, Ohira K, Takao K, Miyakawa T. Transcriptomic evidence for immaturity of the prefrontal cortex in patients with schizophrenia. Mol Brain. 2014;7:41.

23. Schwede M, Nagpal S, Gandal MJ, Parikshak NN, Mirnics K, Geschwind DH, et al. Strong correlation of downregulated genes related to synaptic transmission and mitochondria in post-mortem autism cerebral cortex. J Neurodev Disord. 2018;10:18.

24. Ginsberg MR, Rubin RA, Falcone T, Ting AH, Natowicz MR. Brain transcriptional and epigenetic associations with autism. PLoS One. 2012;7:e44736.

23

# DISCUSSION

ASD is a highly heritable neurodevelopmental disorder. In recent years, hundreds of candidate genes have been related to this condition, although it is still unclear how genetic factors contribute to ASD. My doctoral thesis has focused on i) investigating the contribution of common and rare variants in candidate genes to ASD and comorbidities and ii) functionally characterizing the effect of loss-of-function of these candidate genes using animal models. Below, I will discuss the advantages and limitations of the methods and models used, and how the results obtained contribute to a better understanding of the genetics of neurodevelopmental and psychiatric disorders.

# 1. The autistic spectrum: The problem of the sample

An important aspect when studying the genetic basis of a complex disorder is the appropriate definition of the case and control samples. It is essential to properly select the individuals to be included in order to accurately identify causal variants or genes specifically related to the disorder. In this thesis, we have performed an association study in a European sample, including Spanish, Dutch and German ASD patients and controls, followed by mutational screenings in a subpopulation of this sample. Also, we have used publicly available data from GWAS, WES and transcriptomic studies to analyse the contribution of rare and common variants in candidate genes to several neurodevelopmental and psychiatric disorders. In the case of ASD, it is relevant to discuss the limitations and implications of the current definition of both the case and control samples as it can have an impact on the results obtained in genetic studies.

## 1.1. Defining ASD patients

In mental disorders, no specific biological markers are available to perform the diagnosis. Therefore, patient diagnosis is mainly based on subjective data collected in a series of pre-defined interviews and assessments, usually based on the Diagnostic and Statistic Manual of Mental Disorders (DSM) criteria. In the case of ASD, the two core elements of children's diagnosis are developmental history, normally obtained from parents, and the observation of the child's interaction during a series of assessments, being the most commonly used diagnostic instruments ADI-R [23] and ADOS-2 [24]. Nevertheless, current diagnostic criteria are controversial in many aspects that need to be discussed.

First, diagnostic criteria for ASD have changed over the years. The previous DSM-IV and DSM-IV-TR editions subcategorized autism into five conditions with distinct features [8,9]. Also, the

International Statistical Classification of Diseases and Related Health Problems 10th revision (ICD-10) considered not only one but five autistic subcategories [314]. In this thesis, ASD individuals from the Spanish sample used for the association study and mutational screenings were diagnosed following DSM-IV-TR criteria and, therefore, present this subcategorization: autistic disorder, pervasive developmental disorder not otherwise specified (PDD-NOS) or Asperger's disorder. In addition, the GWAS data we used from Grove et al. [174] included an ASD sample from the integrative Psychiatric Research consortium (iPSYCH) in which cases were selected according to ICD-10 and present the following categorization: childhood autism (ICD-10 code F84.0), atypical autism (84.1), Asperger's syndrome (F84.5), other pervasive developmental disorders (F84.8), and pervasive developmental disorder, unspecified (F84.9). In our association study we did not consider these subtypes and pooled all patients together in the case sample. This was based on previous studies that had pointed to a lack of clear borders between subtypes and the subjective categorization of the patient strongly dependent on the clinician performing the diagnosis [18]. Similarly, the summary statistics we obtained from Grove et al. [174] did not consider autistic subtypes and all the cases were analysed as a "general" ASD sample.

Recent research has also highlighted the heterogeneity in the developmental trajectories of autistic children: there is a high variety in the age of onset of first symptoms and these symptoms can improve or get worse over time depending on the patient [38,315,316]. Therefore, clinical subcategorization would also depend on the life time point in which diagnosis is performed. Finally, autistic patients frequently meet criteria for other mental disorders, such as ADHD, ID or epilepsy. Indeed, in the Spanish sample used in this thesis, 27% of ASD patients present ID and 25% present comorbid disorders such as epilepsy, anxiety or meet criteria for ADHD - as in DSM-IV-TR and previous editions, ADHD could not be diagnosed along with ASD. In the case of the ASD GWAS from Grove et al., 14.3% of ASD cases present ID and no information is available about other medical conditions. The high frequency of co-occurring disorders contributes to autistic spectrum heterogeneity and questions the utility of studying specific mental disorders isolated from others. In the last and current editions of DSM (DSM-5) [10] and ICD (ICD-11) [317], ASD subcategories were removed and the definition and diagnostic criteria were modified to avoid subjectivity bias and to facilitate the diagnosis. In the DSM-5 edition, symptomatic domains are reduced to two, instead of three, and co-diagnosis with other neurodevelopmental, mental or behavioural disorders is possible.

These continuous changes in the definition of autism reflect the difficulty in establishing consensual diagnostic criteria in a disorder with a high clinical heterogeneity [18]. Indeed, current ASD definition encompass a wide spectrum of individuals that fulfil symptomatic criteria

but with strong phenotypical differences, ranging from high-functioning individuals to patients with severe intellectual disability. Importantly, this high heterogeneity in ASD patients may reflect a complex and heterogeneous genetic basis and divergent neurobiological mechanisms, making difficult to find a unified pathophysiology and common causal genetic factors underlying this heterogeneous phenotype.

All the aspects discussed here challenge current diagnostic criteria, reveal the heterogeneity among ASD patients, and antagonize with the static case-control dichotomy. In an attempt to redefine diagnostic criteria of mental disorder, RDoC (Research Domain Criteria Initiative) was created in 2009 as a research framework to integrate different levels of information in a matrix, in which not only symptoms are considered, but also genetics, physiological and behavioural data [318]. The aim of the RDoC matrix is to evolve and be modified based on new research findings in order to help mental health screening tools, diagnostic systems and treatments.

## 1.2. Defining controls

Properly defining controls is highly important to allow comparisons with the case group. The control group needs to have a similar ethnicity component and equivalent sex proportions to avoid detecting differences between samples that are not due to the variants or genes studied but to other confounding factors. In ASD, it is especially important that the sex proportions are similar due to the strong difference in the prevalence of the disorder in males and females (male:female ratio, 4:1) [15], a circumstance that is the rule rather than the exception in psychiatric disorders. Usually, a random group of healthy individuals from the general population is used, with the same ethnicity and sex proportions than controls, and we did so in our association study. In the GWAS performed by Grove et al. [174], controls have two origins: a sample from the iPSYCH in which control individuals come from the same population cohort than cases and include children without an ASD diagnosis, and a sample from the Psychiatric Genomics Consortium (PGC) consisting of parent-proband trios in which pseudo-controls were obtained as simulated individuals derived from the untransmitted parental alleles. Sex proportions are not that similar in the case of Grove et al., with 49.3% of females in the control sample and 21.6% of females in the ASD sample.

A recent viewpoint suggests ASD as a continuum phenotype with a normal distribution of autistic traits in the general population, in which autism diagnosis is at one tail of the distribution [196,197]. Following this model, a control population would contain individuals that may present under-threshold autistic traits and may share genetic variants underlying these autistic

phenotypes with the ASD population (Figure 19) [198]. Indeed, Grove et al. found a genetic correlation between ASD cases and a non-ASD population presenting social communication difficulties [174]. Therefore, comparing ASD patients with individuals considered controls but presenting slight autistic traits could lead to false negative results, compromising the investigation of the underlying mechanisms of the disorder. Even though this has to be considered as a limitation, especially when studying the genetic contribution of variants with small effect sizes, characterizing autistic traits in a control population is challenging and would difficult recruitment and data recovery.
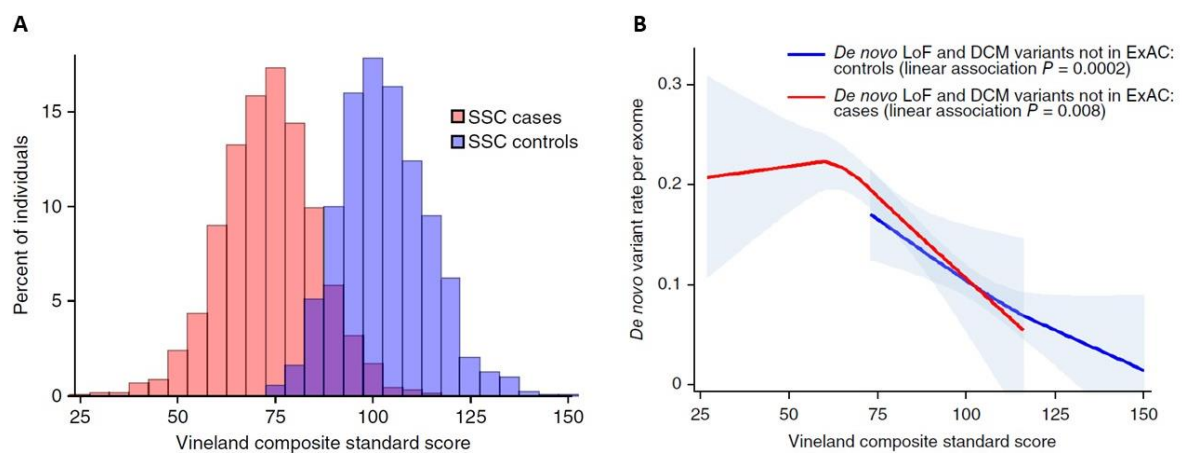


**Figure 19. Overlap between ASD cases and controls suggest a continuum influence of genetic variants. A)** The distribution of Vineland adaptive behaviour scores shows an overlap between cases and control samples from the Simons Simplex Collection (SCC). The Vineland Adaptive Behaviour Scales (Vineland) [319] calculates a score based on parent-rated abilities in social, communication and daily living skills, and normalizes these abilities to a mean of 100 and a standard deviation of 15 in the population (corrected for ages). The SCC population is formed by ASD patients (cases) and unaffected siblings (controls). **B)** *De novo* variation found in case and control samples is associated with the Vineland adaptive behaviour score, which suggests that *de novo* variation influences a continuum of phenotypes in ASD cases and controls. Natural (loess) association is represented and p-values are calculated from Poisson regression. Shaded grey regions represent 95% confidence intervals for association. Adapted from Robinson et al., 2016 [198].

## 2.     Genetic studies in ASD and comorbidities

The current genetic model of ASD postulates an additive contribution of both common and rare variants to the aetiology of the disorder. Different computational and functional approaches exist to characterize the effect of variation in candidate genes to a given disorder. In what follows, I will comment on the advantages and limitations of the main methodologies used in

this thesis to study the contribution of common and rare variants in candidate genes to ASD and comorbid disorders and discuss the results obtained.

## 2.1. Studying common variation

Both large- and small-scale case-control association studies are used to identify genetic risk factors for complex disorders by comparing allele frequency between a group of cases and a group of unrelated controls. These studies investigate the contribution to the disease of polymorphic variants situated in a specific region of interest or in the whole genome. Only common variants (SNPs) are considered (usually MAF ≥ 0.05 or 0.01), and rare variants with strong phenotypic effects are discarded from the study. Initially, association studies made the focus on one or more candidate genes previously related to the phenotype in clinical, pharmacological, molecular or genetic studies. Later, with the emergence of high-throughput genotyping technologies, studies began to analyse SNPs situated along the whole genome in hypothesis-free GWASs, which has permitted to identify new genes and functional pathways without being biased from previous -and limited- knowledge.

In this thesis, we have performed a case-control association analysis using a sample that includes ASD patients and controls to investigate the contribution of common variants in the 14-3-3 gene family to autism (Article 1). Also, we have used publicly available GWAS data to investigate the contribution of common variants in 14-3-3 gene family, *RBFOX1* and *BEX3* to psychiatric disorders (Articles 1 and 3).

### 2.1.1. Testing the contribution of 14-3-3 to ASD through an association study

Association studies have several limitations inherent to the methodology used. Two main types of errors can be distinguished: false positives (type I error), that arise when we detect associations that are not real, and false negatives (type II error), that arise when we are not able to detect real associations. To minimize these types of errors in our population-based association analysis we need to consider several aspects, reviewed in the following paragraphs.

Statistical power of the sample

The statistical power of an analysis is the probability of detecting a true association when it exists. Several factors influence the estimated statistical power, such as disease prevalence, linkage disequilibrium (LD) between the marker and the causal variant, allelic frequencies and risk conferred by the allele (measured with odds-ratios, OR). However, the most important one

is the size of the sample [320]. In our association study (Article 1), the sample used was limited to 727 ASD patients and 714 controls and the analysis failed to identify SNPs associated with the disorder. The sample size has an inverse correlation with the phenotypical impact of the variants detected [320]. Thus, in our analysis an increased sample size would be necessary to detect variants with a small effect size in ASD.

<u>Homogeneity of the sample</u>

When designing an association study, it is important that the phenotype is homogeneous in the whole sample to avoid a reduction of statistical power [320]. Our ASD sample includes only patients from European populations (Spanish, Dutch and German) and with European ancestry. Patients were all diagnosed using ADI-R, and ADOS when possible, and following DSM-IV-TR criteria in order to have a sample as homogenous as possible. Even though DSM-IV-TR edition considered ASD subtypes, the majority of association studies include several of these subtypes in a "spectrum" sample due to the unclear limits between these categories [168,172]. For this reason, our ASD sample included patients diagnosed with either autistic disorder, PDD-NOS or Asperger's disorder. Nevertheless, as discussed before, DSM criteria for ASD comprise a wide heterogeneous group of individuals. Disease heterogeneity reduces statistical power of the analysis because each locus would only determine susceptibility for a subset of the cases, which makes difficult to identify variants associated with the disorder. Additionally, some individuals present other comorbidities that add variability to the phenotype. However, information about comorbid diseases is sometimes not available for all the cases included in the study, which prevents to consider this factor in the analysis.

<u>Defining the control sample</u>

As mentioned before, it is essential to properly define the control sample when comparing it to a group of patients. The control sample needs to have similar characteristics to the case sample in terms of ethnicity, sex proportions and age, so that differences in genetic variation could only be attributed to the disorder and not to other confounding parameters. Also, the number of controls is usually equal to or higher than the number of cases in order to gain statistical power and avoid false positives [320]. In our study, both cases and control groups have European ancestry, a comparable number of individuals (727 cases and 714 controls) and a similar proportion between sexes (84.7% males in cases and 84.5% males in controls). The autistic phenotype has not been specifically excluded from the control sample, but the impact of the contamination of the control sample with affected individuals is minimal, as the prevalence of ASD is only around 1%. Finally, we need to consider population stratification, as geographic,

social or cultural differences between cases and controls may lead to false positives. In our case, case-control number and sex proportions are equal not only in the whole sample but also in each of the three populations included. Nevertheless, these stratification aspects normally have a low impact in the analysis if cases and controls are homogeneous in ethnicity, which is the case in our sample.

## Selection of variants

In hypothesis-driven association studies, variants are selected and genotyped based on previous knowledge or data that point to a contribution of these variants to the disorder. These selected variants are usually located in candidate genes or suspected to have a functional impact on the disorder. In our study, the aim was to investigate the contribution of common variants in the 14-3-3 gene family to ASD. Therefore, 42 tag SNPs with a described minor allele frequency (MAF) ≥ 0.05, a predefined threshold of common variation, were selected to cover the seven 14-3-3 genes. Thirty-four of these SNPs survived quality controls, showed a MAF > 0.05 in our sample, and were then examined for association with ASD.

## Genotyping errors

In some occasions, genotyping errors may occur, due to a low quantity or quality of the DNA samples or data manipulation, and will affect our results. For this reason, it is important to include duplicates and negative controls in the samples to be genotyped. In addition, the genotyping rate is calculated for each individual and only individuals over a certain percentage are considered in the analysis [320]. In our association study, duplicates were included as controls for genotyping quality. Individuals with genotyping rate below 90% were removed from the study, setting the final sample to 1405 individuals (713 cases and 692 controls). Quality controls are also performed to discard SNPs that will not be included in the analysis. From the 42 SNPs initially genotyped, four assays failed, one SNP was excluded for quality reasons and another one for being monomorphic. The call rate for the 36 remaining SNPs, which means the proportion of individuals in the study for which the corresponding SNP information is available, was 99.2% and all were in Hardy–Weinberg equilibrium (threshold set at $p < 0.01$ in controls). Hardy-Weinberg equilibrium is calculated by comparing the genotype and allele frequency distribution, and is a useful quality control tool, as deviations in this equilibrium may be due to genotyping errors [321].

## Multiple testing correction

In association studies, multiple genetic variants are often analysed together and a high number of comparisons are performed. The probability that one statistical test is significant increases

with the number of statistical tests performed. Hence, performing a high number of statistical tests increases the odds for a false positive result and it is important to correct for multiple testing and restrict the significance threshold according to the number of tests performed. The most commonly used multiple-testing correction is Bonferroni, in which significance threshold directly depends on the number of tests performed/variants studied/ models considered (n), and is calculated as follows: p-value = 0.05/n. The Bonferroni correction assumes independence between tests although, in some instances, SNPs are in linkage disequilibrium and therefore they are not independent [321]. Bonferroni is very conservative and the probability of obtaining false negatives is high, which has prompted consideration of other multiple testing correction methods, such as False Discovery Rate (FDR) [321].

In our case, the Bonferroni correction was used and calculated according to the final number of SNPs analysed (p = 0.05/34), but no variant overpassed the suggestive threshold. The main reasons for this study to fail in identifying SNPs associated with ASD may be the small sample size and the low number of SNPs considered, which decreases the statistical power and makes difficult to identify variants with small effects. Indeed, to avoid these under-power issues, and to identify variants across the whole genome, GWAS are becoming increasingly popular in psychiatric research.

### 2.1.2. Genome-wide association studies (GWAS) in mental disorders

GWAS are hypothesis-free association studies with a high coverage and sensitivity that aim to identify common variants with small effect sizes associated with a given disorder. Compared to "traditional" hypothesis-driven association studies, GWAS main advantages are not being biased by previous knowledge and that the typically large sample sizes allow detection of variants with small effect sizes that were previously undetectable [322]. In the last years, big consortia, such as the PGC, have been created to share data from different cohorts and therefore to increase the statistical power of GWAS analyses. In this thesis, we have used publicly available GWAS data from the PGC but also other big cohorts to analyse the contribution of common variants located in our candidate genes, 14-3-3 (Article 1), *RBFOX1* (Article 3) and *BEX3*, to ASD and other neurodevelopmental and psychiatric disorders.

<u>GWAS samples</u>

As in small-scale association studies, in GWAS the patients sample has to be as homogeneous as possible and the case and control groups need to be similar to minimize confounding factors that may affect results [323]. In our GWAS analyses, we have only considered summary statistics

including cases and controls with European ancestry. In the case of ASD it is also important to consider the phenotypical heterogeneity of the spectrum. It is possible that disease heterogeneity reflects differences in the underlying genetic factors and that common variants do not contribute equally to the whole spectrum. Indeed, Grove et al. showed differences in heritability and a strong heterogeneity of genetic overlap with other traits among ASD sample subsets. This work also showed that the polygenic contribution of common variants seem to be more important in high functioning ASD patients than in patients with comorbid ID [174]. In line with that, a recent small-scale GWAS in ASD showed that dividing cases into subgroups based on phenotype clustering successfully identified significant associations that were not found without clustering [324]. Also, the presence of comorbidities in some ASD patients suggests possible differences in the genetic architecture of the disorder between individuals. Indeed, a recent preprint work from Mattheisen et al. [325] reveals specific patterns of polygenic architecture that characterize comorbid ASD and ADHD and that are different from the ASD-only cases. In particular, they show that comorbid ASD and ADHD cases present different patterns of genetic association with other traits when compared to the ASD-only subgroup.

Selection of variants

In GWAS, thousands of polymorphic variants (usually MAF ≥ 0.01) distributed evenly throughout the genome are automatically genotyped. Only variants and individuals with a genotyping rate over 80-90% are included in the analysis. In addition, many GWAS are indeed meta-analyses that often include data from studies that have used different array platforms, and therefore a different set of SNPs. To solve this issue, an imputation step is performed in which genetic markers that have not been directly genotyped are inferred, based on the LD of a reference sample [322,323]. In our studies, all the GWAS summary statistics data used were checked and filtered to fulfil these quality criteria: MAF ≥ 0.01 and info-score for imputation quality ≥ 0.6.

Multiple testing correction

A high number of polymorphisms is tested and, therefore, a huge sample and a conservative threshold are necessary to avoid false positives. GWAS used a predefined genome-wide association significance threshold of p-value < 0.05/1e-06 = 5e-08, based on the Bonferroni correction for a million of independent tests, and a  pre-defined "suggestive" association significance threshold: p-value < 1e-05 [322,323,326]. We have applied both these predefined significance thresholds in our SNP-based analysis.

Association studies using available GWAS summary statistics (SS)

Using filtered summary statistics from GWAS, we explored the individual contribution of common variants in *RBFOX1* and flanking regions (between 3000-20000 SNPs, depending on the SS) to several psychiatric disorders using a SNP-based analysis. We found genome-wide significant associations between several single SNPs in *RBFOX1* and major depressive disorder (MDD), risk taking (RT) and the cross-disorder meta-analysis (CD-MA). These disorders are the ones with a higher sample size (MDD: 135,458 Ca + 344,901 Co; RT: 975,353 individuals; 232,964 Ca + 494,162 Co), which provides more statistical power to detect variants with a small effect size. Indeed, the next GWAS with more individuals is schizophrenia (SCZ, 67,280 Ca + 86,912 Co), for which *RBFOX1* shows a nominal association. This suggests that increasing sample size in GWASs for the other disorders explored may lead to significant associations that have not been detected yet due to the current under-power of the analyses. A clear example of the importance of sample size in GWAS is ASD. The first GWASs performed in ASD did not have enough statistical power due to small sample sizes [168,169,172,173], but the last well-powered ASD GWAS, almost tripling the previous largest discovery sample with a cohort of more than 18,000 cases, was finally successful to report five loci with a small effect size associated with ASD [174]. Although this last study provides an important step towards understanding common variation in ASD, only five loci were identified as significantly associated with the disorder, and we are still far from gaining enough statistical power to identify a higher number of robust associations. Other psychiatric disorders that have achieved larger sample sizes, such as SCZ, have already surpassed 100 genome-wide significant risk loci [327].

Importantly, an issue when analysing GWAS results is to determine which are the causal variants among the ones identified. If the associated SNP is in a coding region and involves an amino acid change, the probability that it is causal is high. However, the majority of the detected tag SNPs are usually in non-coding regions and far from known genes. In our case, all SNPs detected in *RBFOX1* significantly associated with a disorder are located in non-coding regions. These SNPs may have an impact in the regulation of the expression of genes situated in *cis* or *trans*, or simply mark the gene or genomic region and, indeed, other variants in LD with them would then be the causal ones [322].

We also performed a gene-based analysis using MAGMA [328] to explore the contribution of common variants in *RBFOX1* and in all members of the 14-3-3 gene family to several psychiatric disorders. MAGMA analysis consists of three steps: first, an annotation step to map SNPs onto genes using the 1,000 Genomes European Panel; second, a gene analysis step in which SNPs located within the transcribed region of the gene are considered to compute gene-wide p-

values; and finally, an optional gene-set analysis in which a single p-value is calculated for a group of selected genes.

We found nominal associations for four out of the seven 14-3-3 genes with several psychiatric disorders and associations with gene-wide significance for *YWHAE* with SCZ and CD-MA. We also performed a gene-set analysis including the seven genes of the 14-3-3 family but only a nominal association was found between the whole 14-3-3 gene family and SCZ. Although almost all the associations we found are nominal, which may be due to the lack of statistical power of the GWAS summary statistics used, these results are suggestive of a contribution to SCZ of common variants in the 14-3-3 gene family, and particularly in *YWHAE*. Indeed, a polymorphism (rs28365859) in *YWHAE* was previously associated with SCZ [211] and correlates with differences in brain morphology in SCZ patients [329,330].

*RBFOX1* showed nominal associations with eight out of eleven several psychiatric conditions, obtaining gene-wide significance for MDD, risk tolerance, SCZ, and the CD-MA. Moreover, significant genes of the gene-based analysis for these four conditions were enriched in *RBFOX1* target genes (using the list of the 2499 genes targeted by the RBFOX1 protein previously described by Lee [233]). These results suggest a pleiotropic contribution of both the *RBFOX1* gene and some of its targets to psychiatric conditions. Again, in all our analyses, significance thresholds are only overcome for the disorders with a larger sample size, and significantly associated variants were not identified in the analyses with smaller samples.

Finally, we aimed to analyse the contribution of common variation in *BEX3* to psychiatric disorders but we experienced several issues. The majority of GWAS focus on autosomal variants only, and the sex chromosomes are not considered due to several technical problems: lower coverage, complications in genotype calling, analytical specificities, imputation and selection of test statistics [331]. We asked for the GWAS summary statistics of the ASD and SCZ GWAS meta-analyses from the PGC [173,327] including chromosome X information, but no variant surpassed the suggestive threshold. As mentioned above, X chromosome genotyping is usually not well defined and has a low coverage, which was the case here. Also, it is important to consider X hemizygosity in males and X inactivation in females, which makes the analysis of the impact of genetic variants in the phenotypes more complicated [331]. All these factors had an influence on the the failure to identify *BEX3* SNPs associated with ASD and SCZ.

The problem of the "missing" heritability

ASD and several comorbid disorders are highly heritable, as demonstrated in early family and twin studies. In the case of ASD, according to a recent meta-analysis, inherited genetic factors

account for around 64-91% of ASD liability [104]. However, genetic studies are still far from identifying all the specific genetic causes underlying this high heritability, and estimates vary across studies. ASD heritability would be defined by common and rare variation, however, only a small percentage of this expected heritability can be currently explained (Figure 20). Although common variation is predicted to account for a high percentage of ASD heritability, in the most powered GWAS study to date only few genetic *loci* were associated with ASD and the SNP-based heritability was calculated to be only 12% [174]. We should also consider that other common variants (CNVs or *indels*) may also explain part of the heritability of ASD, although their contribution has not been clearly quantified. The emergence of GWAS was thought to help explain the missing heritability in complex psychiatric and neurodevelopmental disorders, however, we are still far from explaining the genetic factors underlying the high heritability of these disorders. In future research, GWAS with larger samples or phenotype clustering of cases may help identify common variants and contribute to a better understanding of this missing heritability in ASD and comorbid disorders.
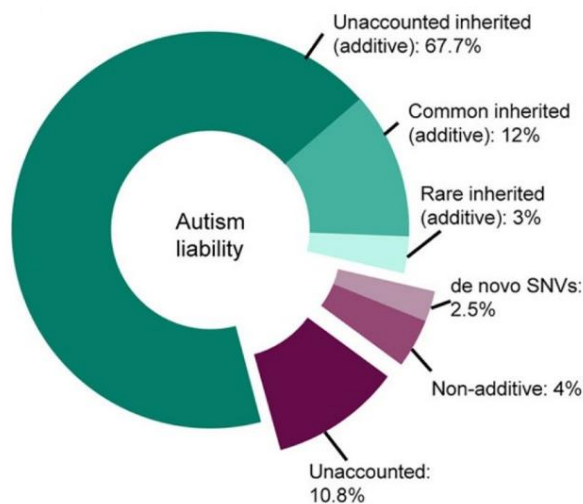


**Figure 20**. **Relative contribution of genetic factors to ASD liability**. Although estimates vary across studies, here we tentatively represent the contribution of different genetic factors to ASD aetiology. Total heritability (82.7%, shades of green) has been estimated using familial recurrence from Bai et al. [332] and includes: common inherited variants (linkage disequilibrium score regression (LDSR)-based SNP heritability from Grove et al., [174], 12%), rare inherited heritability ([111], 3%) and the currently unexplained additive heritability (67.7%). The remaining ASD liability (shades of purple) would include *de novo* variants, non-additive variation [205] and currently unaccounted factors. Here, *de novo* variation estimates include missense and protein-truncating variants [205] and variation in non-genic regions [333] but does not include *de novo* structural variants and tandem repeats, thus it may be lower-estimated. Adapted from Havdhal et al., 2021 [260].

### 2.1.3. Imaging study of a common variant in *RBFOX1*

Functional magnetic resonance imaging (fMRI) is a useful tool to investigate the effect of genetic variants in brain circuitry related to psychiatric traits. In this thesis, the effect of a common variant (rs6500744; aggression risk allele: C) situated in the first intron of the *RBFOX1* gene in brain circuitry implicated in emotion processing, fear conditioning and executive functioning

was explored using fMRI. Specific rare variants cannot be investigated using this technique with enough statistical power, and the selection of this specific common variant for the study was based on a previous association study that related the C-allele risk to aggressive behaviour, more precisely to conduct disorder symptoms [334].

In a first approach with healthy volunteers, C-allele carriers showed an increased activity in the dorsal anterior cingulate cortex (ACC) in response to fearful and angry faces, and a reduced left dorsolateral prefrontal cortex (DLPFC) activity during tasks requiring cognitive control, compared to T/T carriers. Additionally, in a group of patients with panic disorder and agoraphobia, C/C carriers presented an increased fear expression and activation of ACC after fear conditioning and a reduced activation in the dorsal ACC and stronger fear reduction after the extinction training in a conditioning paradigm. Interestingly, these patients had significantly increased depression and anxiety scores and increased levels of RBFOX1 protein. Finally, in a larger group of subjects with panic disorder and agoraphobia, the C-allele was found to be dose-dependently associated with the frequency of avoidance behaviour as shown in a behavioural avoidance task (BAT) assay [335], results that were in line with the patients' clinical ratings of everyday life avoidance behaviour.

These results suggest that C-allele carriers present an increased reactivity to emotional *stimuli*, given the role of ACC in integrating cognition with emotion [336–339], and in fear appraisal and expression [340–342]. C-allele carriers also show an altered activation of DLPFC, suggesting a deficient processing during cognitive and impulse control [343,344]. These differences in brain activation when processing emotion and during executive functioning may contribute to an increased risk for psychiatric disorders characterized by increased emotional reactivity (MDD), impaired impulse control (ADHD, ASD, addictions…) and aggression [345,346]. Importantly, these psychiatric conditions had already been associated with *RBFOX1* in terms of common or rare variation. Finally, differences in the level of RBFOX1 protein expression in individuals with the CC-genotype with respect to the T-allele carriers also suggest a contribution of *RBFOX1* to the phenotypical alterations observed.

## 2.2. Studying rare variation

Next-generation sequencing at exome- or genome-wide scale, chromosomal microarray studies but also sequencing of target genes have helped to identify a high number of inherited and *de novo* rare variants that contribute to ASD and other mental disorders. In ASD, variants with large effect size are rare and contribute mainly to the most severe phenotypes, usually with comorbid

ID (Figure 21) [181]. Although the collective contribution of pathogenic rare variants seems to be more important in neurodevelopmental disorders, rare variation has also been identified in other psychiatric conditions with an early onset such as ADHD or schizophrenia, and, less frequently, in patients with mood disorders [347]. Indeed, Satterstrom et al. found a similar burden of rare protein-truncating variants in ASD and ADHD patients in an exome sequencing study of 13,000 individuals [348].
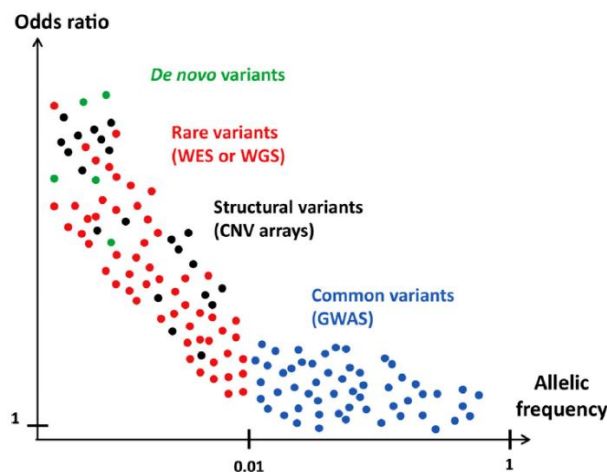


**Figure 21. Contribution of common and rare variants to ASD.** Toma, 2020 [181].

### 2.2.1. Whole-exome and whole-genome sequencing: identification of point mutations and CNVs

In the last years, a huge number of sequencing studies have investigated the contribution of rare variants to mental disorders and different databases have been created to gather these increasing amounts of data and to make them publicly available. Some examples of these datasets are the Genome Aggregation Database (GnoMAD), the Geisinger Developmental Brain Disorders Database, ClinVar or DECIPHER. Also, the Broad Institute website offers data from case-control studies of psychiatric diseases, and the database of Genotypes and Phenotypes (dbGaP) includes data from a huge number of genetic studies linking variation in the genome to specific phenotypes and disorders. The first genetic studies in ASD focused on rare variants as they were initially considered the main contributors to ASD genetic liability. For this reason, the number of studies on rare variation are significantly higher in ASD compared to other mental conditions. In this thesis, we have used data from previous WES and WGS published studies and databases to analyse the contribution to ASD and comorbidities of rare variation in the 14-3-3 gene family (Article 1), *RBFOX1* (Article 3) and *BEX3* (Article 5).

Rare variants in the 14-3-3 gene family

We investigated the impact of both rare (MAF < 0.01) and ultra-rare (MAF < 0.0001) variants reported in the 14-3-3 genes to ASD and SCZ using information from large datasets including

information from family-based and case-control studies, as well as our own European sample for the ultra-rare variants (URVs) analysis.

Following the identification of rare variants in a sequencing study, it is important to estimate their pathogenicity using predictive tools (and, eventually, functional studies) in order to evaluate their possible contribution to the disorder. The selection of URVs to be analysed was based on their predicted pathogenicity using the Variant Effect Predictor annotation tool software, which provides information about the variant type and it calculates the probability of a variant to be damaging using both SIFT and PolyPhen-2. In addition, to analyse the contribution of rare variants to specific disorders, burden tests are usually performed. In our studies, we used burden tests to compare rare variation frequency in ASD or SCZ patients compared to controls. We found that only *YWHAZ* reached significance for a higher number of rare SNVs in SCZ. In the URVs analysis, given the relatively limited number of genetic variants found, data for the whole gene family were combined. URVs in the 14-3-3 genes were found to be three times more frequent in ASD patients than in controls and a significant burden of URVs was found in ASD patients when compared to controls – mainly due to a splice variant in *YWHAE* found in four unrelated ASD probands. For SCZ, no significant burden was detected.

Our results suggest that URVs are more frequent in ASD patients than in controls. Ultra-rare variants are usually more pathogenic, as natural selection limits their frequency, and they have been usually described in patients with neurodevelopmental disorders and severe phenotypes [181]. In our study, the same URV was found in *YWHAE* in four unrelated patients. This gene plays an important role in neuronal morphogenesis [208,349], and knockout mouse studies have shown that the loss of *YWHAE* function leads to defects in brain development, neuronal migration and behaviour, showing defects in working memory, anxiety-like behaviour and hyperactivity [211,221,223]. Indeed, microduplications of this gene were reported in ASD patients and microdeletions involving both *YWHAE* and *PAFAH1B1* cause Miller-Dieker syndrome, a neurodevelopmental disorder with severe cognitive impairments [350–354].

We also found rare SNVs to be more frequent in *YWHAZ* in SCZ cases than in controls. Interestingly, a mouse model with *YWHAZ* deficiency presents hippocampal alterations and behavioural and cognitive problems, including disrupted sensorimotor gating, that are related to psychiatric disorders and more precisely to SCZ [220,222]. Moreover, 14-3-3ζ (encoded by *YWHAZ*) interacts with the Disrupted-in-Schizophrenia 1 (DISC1) protein, an important schizophrenia risk factor [220,222]. These findings suggest that damaging variants in *YWHAZ* may affect protein function and interaction leading to a schizophrenic phenotype. Finally, some 14-3-3 proteins, are especially enriched in synapses, where they interact with pre-synaptic

proteins, regulate transmission and plasticity [355]. Although rare variants are less frequently described in SCZ compared to ASD, exome sequencing studies have found *de novo* protein-disrupting variants to be enriched in synaptic genes, genes involved in glutamate signalling and in cytoskeletal pathways [356,357].

Copy number variants in *RBFOX1*

An exhaustive literature search was performed using PubMed to gather all available information about CNVs spanning *RBFOX1* in patients and controls when possible. We identified CNVs in patients diagnosed with ADHD, ASD, BD, OCD, SCZ, TS or multiple mental disorders/traits. The vast majority of the reported CNVs were found in ASD patients (112 CNVs), but also in SCZ patients (24 CNVs), and spanned *RBFOX1* specifically. These results are in line with the higher contribution of rare variation to neurodevelopmental disorders compared to other psychiatric disorders, in which common variation seems to have a higher contribution. Indeed, they point to a higher contribution of *RBFOX1* rare variants to ASD and SCZ than to other psychiatric disorders. Nevertheless, it is important to consider that CNVs studies are more commonly performed in ASD patients than in other disorders due to previous findings suggesting a strong implication of rare variants in neurodevelopmental disorders, which may bias our results.

CNVs were also found in some control individuals, but *RBFOX1* CNVs are 2.3 times more frequent in cases than controls, and 5 times more frequent in ASD patients than in controls. Burden tests could only be performed for few studies, as in the majority of them CNVs in patients were not investigated. More information about CNVs in controls and in patients diagnosed with psychiatric disorders would be necessary to better assess the contribution of CNVs to these disorders.

Rare variants in the *BEX/TCEAL* gene family

We first investigated the presence of rare variants in the *BEX3* gene by exploring several WES datasets but we could not find any SNV reported in ASD or SCZ patients surviving quality filters, nor CNVs spanning *BEX3* specifically, although this could be due to the small size of the gene or to a negative selection of variants that may have a strong effect. However, we found several reported CNVs spanning various *BEX/TCEAL* genes in patients with severe neurodevelopmental impairments. First, a duplication spanning *BEX3*, *TCEAL4*, *TCEAL9*, and *RAB40A* genes was reported in a patient with ASD (Decipher database, ID: 290829). Also, an inherited duplication spanning a region containing the whole *BEX* gene family and some *TCEAL* genes was reported in a patient with infantile spasms, a specific type of epileptic encephalopathy associated with severe developmental disabilities [358]. Finally, several Xq22 deletions that affect *BEX3* and

some other *BEX/TCEAL* genes were reported in patients diagnosed with severe early-onset neurological disease trait (EONDT) or general severe neurological impairments [250–252]. These reported CNVs point to an important role of *BEX3*, but also other *BEX/TCEAL* genes, in neurodevelopment and associated disorders. We hypothesize that *BEX/TCEAL* genes dysfunction may contribute to severe neurodevelopmental phenotypes.

### 2.2.2. Mutational screening of *BEX3* and the 14-3-3 gene family

Mutational screenings can be performed in patient samples to specifically explore the presence of rare variants in candidate genes. In this thesis, we used our European ASD sample composed of 288 patients (182 Spanish, 94 Dutch, 12 Germans) to investigate the presence of rare variants in the 14-3-3 genes and in *BEX3*.

*BEX3* is a relatively small single-exon gene and, therefore, the mutational screening was performed by Sanger sequencing: including the coding region, 5' upstream sequence and 3'UTR (chrX:103,375,910-103,378,077). However, no rare variants were found in our ASD patients, which is in line with the literature search results discussed above. In our 14-3-3 screening, high-throughput sequencing was performed using the Ion Torrent platform and 57 tagged-primer pairs. Sequencing covered 96.3% of all coding exons across the 14-3-3 gene family, including the splice sites and part of the 5' and 3' untranslated regions (UTR). Nine rare variants were found and subsequently validated by Sanger sequencing. Pathogenicity was predicted using SIFT and Polyphen2 tools: two variants were predicted to be deleterious and were found in *SFN* in the same patient together with a third variant predicted to be benign. All of them were inherited from the same maternal chromosome (Article 1).

### 2.2.3. Functional characterization of rare damaging variants

Following the discovery of rare variants predicted to be damaging, it is important to perform functional studies to characterise the effect of these variants in protein or gene function and, therefore, in the phenotype. In this thesis, we have performed a series of functional *in vitro* studies to characterize the effect of predicted damaging variants in the *YWHAZ* and *SFN* genes and to evaluate their contribution to ASD in these patients (Article 1).

In a previous study of the group, a frameshift variant (c.659-660insT, p.L220Ffs*18) in the *YWHAZ* gene, inherited from a mother with depression, was identified at heterozygosity in two siblings diagnosed with ASD and ADHD [227]. *YWHAZ* encodes 14-3-3 $\zeta$, that forms heterodimers

with 14-3-3ε and exerts its function binding to different proteins such as tyrosine hydroxylase (TH), an enzyme that is essential for DA synthesis [359,360]. Using surface plasmon resonance, the truncating mutation identified in *YWHAZ* was found to prevent the mutated 14-3-3 ζ to bind the phosphorylated serine 19 of TH. 14-3-3 proteins are molecular chaperones that help other proteins to fold properly. Thus, an impaired interaction between 14-3-3 ζ and TH may have implications in dopamine synthesis, a key neurotransmitter involved in diverse cognitive and executive functions and previously related to ASD and SCZ [102,361]. We also explored the capacity of dimerization of the mutated 14-3-3 ζ using Bioluminescence Resonance Energy Transfer (BRET) assays in HEK 293T cells. We found that *YWHAZ* truncating mutation impaired protein dimerization with 14-3-3 ε, which may explain the impossibility to bind to TH, previously described, and probably to other proteins. Therefore, an impaired dimerization of 14-3-3 ζ may have implications in important neurological pathways and lead to the onset of neurodevelopmental disorders.

We also explored the effect of the three mutations (two of them damaging) found in the *SFN* gene, encoding 14-3-3 σ, in an ASD patient. Using a BRET assay, we found that the 14-3-3 σ mutant protein was still able to form homodimers. Even though these results suggest that these rare variants in *SFN* do not impact dimerization, they may affect protein binding. We wanted to explore the effects of these mutations in 14-3-3 σ binding with Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1A (DYRK1A) protein, a previously described interactor [362–365] strongly related with neurodevelopmental disorders and ASD [366,367]. However, DYRK1A and 14-3-3 σ binding is a transient interaction that may only occur in specific moments. Thus, we were not able to detect the interaction between DYRK1A and WT 14-3-3 σ using a BRET assay, and could not explore the effects of these mutations in this protein-protein interaction. Importantly, we should consider that all these functional studies were performed *in vitro* to model a living system. These models can be used as a first approach to analyse the functional effects of specific mutations but present limitations, as they do not represent the real situation in a living organism in which other parameters may modulate or contribute to the observed effects.

## 2.3.  Transcriptomic data

Microarray technology has been widely used to analyse gene expression in specific samples or tissues. More recently, RNA sequencing technology (RNA-seq) emerged as an attractive alternative due to its higher specificity and sensitivity and its ability to detect novel transcripts.

Both approaches allow obtaining information about gene expression in *postmortem* brain samples and analysing possible differences between patients and controls [368,369]. First criticism in the field regarded rapid degradation of RNA and long *postmortem* intervals, factors that would affect RNA quality. However, optimized protocols have been successful in obtaining high quality RNA. Nevertheless, other limitations of *postmortem* samples may have affected our results, such as variability in the interval between death and the collection of the sample, cause of death and agonizing state of the individual. *Postmortem* transcriptomic data give us information about the genetic profile of patients with a given disorder. This presents an advantage when compared to *in vitro* cell culture and animal models, as several studies have pointed to a specific human gene expression profile, different from other species [370,371]. The increased interest in transcriptional profiles of the human brain in different psychiatric disorders led to the creation of huge datasets such as the Gene Expression Omnibus (GEO), a public functional genomics data repository.

In this thesis, we have collected publicly available data from published articles and the GEO database about gene expression in *postmortem* brains of ASD and SCZ patients. We aimed to analyse possible changes in the expression of the 14-3-3 gene family (Article 1), *RBFOX1* (Article 3) and *BEX/TCEAL* gene families (Article 5) in patients with these disorders. We first reviewed all the studies reporting significance differential gene expression (DGE) and, in the case of the *BEX* and *TCEAL* gene families, we also calculated the enrichment of differential gene expression using a hypergeometric test.

We found that several studies had reported a decreased *RBFOX1* expression in cortex in ASD patients and in different brain areas in SCZ patients, although only differences found in three of the studies overcame the 10% false discovery rate (FDR) significance threshold: two studies in ASD and one study in SCZ patients. FDR is a multiple testing correction that calculates the proportion of false positives among all the positive results obtained and readjusts the significance threshold. In the case of 14-3-3 genes, we found a decreased expression of *YWHAB*, *YWHAE*, *YWHAH* and *YWHAZ* in both ASD and SCZ patients and an increased expression of *SFN* in both disorders. Also, *YWHAQ* showed a decreased expression only in ASD patients. All these differences overcame a 10% FDR significance threshold and were not tissue-specific, as they were observed in different brain areas. Finally, we found that all *BEX/TCEAL* genes were downregulated in ASD and SCZ patients in different brain regions, and in many cases this differential gene expression overcame FDR correction. In addition, *BEX* genes were significantly enriched among the differentially expressed genes in most datasets. In summary,

downregulation of all these candidate genes in ASD and SCZ patients suggests their implication in these early-life onset neurodevelopment disorders.

# 3. Animal knockout models of ASD and comorbidities

Animal models offer preclinical tools that help to get more insight into the genetic and molecular mechanisms underlying neurodevelopmental disorders. Although mice have always been the most popular animal models in medical research, in the last years the zebrafish has emerged as an alternative powerful model (for more details, see chapter 3 of the Introduction section). In this thesis we have used mouse and zebrafish models to investigate the role of *YWHAZ*, *RBFOX1* and *BEX3* in ASD and psychiatric disorders using behavioural, molecular and imaging approaches.

## 3.1. *YWHAZ* models

Previous studies had investigated the role of the 14-3-3 genes, and more precisely *YWHAZ*, in neurodevelopment using KO mouse models [220,222,226]. To further explore the contribution of *YWHAZ* to neurodevelopmental and psychiatric disorders, in this thesis we studied the role of *ywhaz* in nervous system development, neuronal activity, neurotransmission and behaviour in zebrafish (Article 2).

First*, ywhaz* KO larvae present differences y neuronal activity and connectivity (results that will be further discussed in the following section) that may be due to impaired neurogenesis, neuronal migration and synaptogenesis caused by *ywhaz* deficiency. We also observed pan-neuronal expression of *ywhaz* during zebrafish developmental stages, which suggests a major role of *ywhaz* in neurodevelopment.

In WT adult fish, *ywhaz* expression was restricted to Purkinje cells in the cerebellum, a region that shows alterations in autistic patients [372]. We found that *ywhaz* KO adults displayed decreased levels of DA and 5-HT in the hindbrain, the area where *ywhaz* is specifically expressed. We also observed that *ywhaz* KO adult fish freeze when exposed to novel *stimuli*, a phenotype that was reversed with fluoxetine and quinpirole, drugs acting on 5-HT and DA neurotransmission. A previous study in *Ywhaz* KO mice with Sv/129 background reported alterations in DA function and SCZ-like behavioural defects such as hyperactivity and a disrupted sensorimotor gating [220]. Another study was performed in *Ywhaz* KO mice with a different

background [222] and some phenotypical differences were reported between the two KO models. However, both of them present learning and memory deficits related to anatomical and synaptic defects found in the hippocampus [220,222]. Importantly, social behaviour was not tested in these models and, therefore, a relation with ASD phenotype could not be assessed.

5-HT plays an important role during brain development by regulating trophic factors and activity-dependent plasticity; it modulates cell division, differentiation, neurite outgrowth and synaptogenesis [373]. It also modulates a wide spectrum of behaviours such as sensory processing, cognition, social interaction and anxiety [97]. Alterations in the 5-HT system may affect neurodevelopment and result in neurological and psychiatric disorders. Indeed, ASD and SCZ patients present alterations in 5-HT levels and in the density of 5-HT receptors across brain areas [97,373,374]. In addition, DA is involved in reward and motivation through the mesocorticolimbic pathway. It has been demonstrated that ASD is characterized by hypoactivation of this reward system [375,376], and patients present deficits for both social and non-social rewards [285,377]. We hypothesize that the decreased level of monoamines observed in *ywhaz* KO could lead to an impaired reward motivation, cognition and sensory processing, which would contribute to the neophobic freezing response of mutants.

All these results suggest that *YWHAZ* plays an important role in neurodevelopment and its impaired function may lead to altered brain activity and connectivity during development with further implications in the adult phenotype. *YWHAZ* deficit seems to affect monoamine signalling and leads to ASD and SCZ-like behaviours, pointing to a role of this gene in the onset of neurodevelopmental disorders. Interestingly, phenotypical differences observed between the two *Ywhaz* KO mice lines suggest a role of the genetic background in modulating the phenotypical outcome of *Ywhaz* loss of function.

## 3.2. *RBFOX1* models

In this thesis we used both mouse and zebrafish models to investigate the *RBFOX1* contribution to psychiatric and neurodevelopmental disorders (Articles 3 and 4).

First, we found that during developmental stages, *rbfox1* presents a pan-neuronal expression in WT zebrafish. These results are in line with previous findings in mouse and cell models that suggest a major role of this gene in neurodevelopment, regulating the splicing and expression of large gene networks implicated in neuronal development and maturation and controlling synaptic transmission [232,245,378].

In addition, to explore the effects of *RBFOX1* deficiency in behaviour, behavioural tests were performed in both mice and zebrafish KO models. A neuron-specific *Rbfox1* KO mouse line was generated to explore the behavioural consequences of a decreased *Rbfox1* expression (Article 2). KO mice showed a pronounced hyperactivity in the open field, light dark box and marble burying tests, thigmotaxis and stereotypic-like behaviour, deficit in the acoustic startle response without impairment in sensorimotor gating, a reduced social interest and lack of aggressive behaviour. Interestingly, both heterozygous KO mice (HET-KO) and homozygous KO mice (HOM-KO) showed impairments in fear acquisition and extinction in the auditory cued fear conditioning test, although HET-KO present a milder phenotype. These results show that *Rbfox1* KO mice present behavioural alterations that can be related to neurodevelopmental disorders such as ASD and ADHD. Interestingly, previous studies in mice showed that specific *Rbfox1* deficiency in the central nervous system leads to impaired neuronal migration, axon extension, dendritic arborisation and synapse network formation [231,233,379]. These previous findings, together with the behavioural alterations we reported in our study, suggest that loss of *Rbfox1* function contributes to the pathophysiology of neurodevelopmental disorders.

Finally, we also reported behavioural alterations in two different *rbfox1* KO zebrafish lines. We found that *rbfox1*[sa15940] mutants present hyperactivity, anxiety-like behaviour, a decreased freezing behaviour, higher levels of aggression and an altered social behaviour, compared to WT TL fish. On the other hand, *rbfox1*[del19] mutants present a similar thigmotaxis behaviour, but stronger alterations in social behaviour and lower levels of hyperactivity than *rbfox1*[sa15940] fish; and, contrary to *rbfox1*[sa15940] fish, *rbfox1*[del19] mutants do not show an aggressive behaviour. Interestingly, both zebrafish KO lines present alterations in behaviour that can be assimilated to ASD and ADHD symptoms, although the genetic background seems to modulate the effect of *rbfox1* deficiency on the behavioural phenotype. In addition, we found that *rbfox1* shows a restricted expression in adult forebrain areas, including telencephalon, hypothalamus and thalamus, regions that play an important role in receiving and sensory processing, stress, and in directing behaviour, especially social behaviour and emotion [380–383]. These results suggest that *rbfox1* deficiency affects forebrain function and contributes to the reported behavioural alterations.

In summary, all *RBFOX1* deficiency models present hyperactivity, thigmotaxis and reduced social interests, which are ASD- and ADHD-like symptoms. Other behavioural alterations are reported to be specific of one or two of the mutant lines, but in all cases the changes observed in both mice and zebrafish models of *RBFOX1* deficiency can be assimilated to phenotypical alterations present in patients with psychiatric conditions such as ADHD or ASD. These results validate all

these KO lines as models for psychiatric disorders and point to *RBFOX1* pleiotropic effects in mental disorders.

## 3.3. *BEX3* model

The involvement of *Bex3* in neurodevelopment and related disorders was also assessed in this thesis. First, *Bex3* was found to be highly and widely expressed during mouse development. Then, morphological, behavioural and molecular alterations were assessed using two *Bex3* KO mutant mice lines: *Bex3*$^{\Delta(24-72)}$ and *Bex3*$^{KO}$ (Article 5). Homozygous mutant mice showed subtle alterations in skull morphology and brain size, with a reduced cortical and cerebellum size and enlarged ventricles. Importantly, enlargement of brain ventricles is a common feature of neurodevelopmental disorders [384–386]. A series of behavioural alterations was also common to both KO lines: repetitive behaviour events, social impairment, impaired sensorimotor gating, and a decreased alternation in the Y-maze test. Cognitive impairment was more severe in the *Bex3*$^{KO}$, as these mice showed a lower performance in the object recognition memory test and the passive avoidance test. All these behaviours can be related to ASD and SCZ traits: repetitive behaviours, social impairment, a search for "sameness", sensorimotor dysfunction and cognitive deficits.

Adult *Bex3*-deficient mice also present a reduced number of cortical and subcortical inhibitory interneurons, particularly in the hippocampal CA2 region of *Bex3*$^{KO}$ mice, and alterations in the excitation/inhibition (E/I) balance due to a strong decrease in inhibitory synaptic transmission in this same region. These morphological and electrophysiological alterations can explain the reported behavioural alterations, as an impaired social interaction and learning has been associated with an altered CA2 hippocampal circuitry [387,388]. Interestingly, *Rbfox1* deficiency in mouse brain leads also to an altered synaptic transmission and E/I imbalance [232,378]. Indeed, E/I imbalance has been reported to contribute to ASD pathophysiology [389,390], and therefore, it might explain ASD-like behaviours in both *Rbfox1*- and *Bex3*-deficient animals. Finally, alterations in the phosphorylation of mTORC1 and mTORC2 readouts were reported, which suggests mTORC2 hyperactivation as a possible molecular consequence of *Bex3* deficiency. BEX3 protein interacts with the TSC1/TSC2 complex, essential for the regulation of the mTOR signalling cascade [391], a pathway described to be disrupted in monogenic neurodevelopmental diseases, such as syndromic ASD [392], but also in psychiatric disorders such as SCZ or MDD [393,394]. Given the involvement of mTOR signalling cascade in neurogenesis and neuronal excitability, we hypothesize that *Bex3* deficiency produces a

dysregulation in mTOR pathway and, consequently, contribute to the observed alterations in E/I balance and behaviour in *Bex3* mutant mice.

## 3.4. Mouse models of comorbid ASD and ADHD

We recently reviewed all transgenic mouse lines reported in the Jackson database presenting both ASD and ADHD-like behaviours (see the corresponding manuscript in Annex). We found 27 lines that present both hyperactivity and ASD-related symptoms such as impaired sociability, repetitive motor behaviours, communication impairments, cognitive rigidity or altered nest building behaviours. In all these lines, a single gene was knocked out and, thus, the behavioural alterations observed were due to the deficiency of this gene. The majority of these genes are listed in the SFARI database as genes implicated in ASD susceptibility, and some of them were previously related to ADHD in patients. Also, the group of 27 genes knocked out in these lines was enriched in genes that participate in regulation of synapse structure and activity, locomotor behaviour and cognition, traits that are disrupted in ASD and ADHD. Interestingly, a recent study performed in mice found that overexpression of different variants of *SHANK2*, one of the genes we reported in our review, lead to different alterations in synaptic function and behaviour, reproducing an ASD-like or an ADHD-like phenotype depending on the overexpressed variant [395]. These results suggest that genetic variation can affect differently gene and isoforms expression and lead to distinct psychiatric-like phenotypes.

In addition, we found that the majority of these transgenic lines show several behavioural alterations that add to those related to ASD or ADHD, with reminiscences to SCZ or MDD. For the other seven lines, only ASD- and ADHD-related behaviours were reported, although this could be due to an incomplete behavioural testing and not to the absence of these behaviours in the KO mice. The co-occurrence of distinct psychiatric-related behaviours reported in these mice lines suggest a pleiotropic contribution of these genes to psychiatric disorders, pleiotropy that is modulated by other genetic and environmental factors.

## 3.5. Benefits and limitations of animal KO models

Although KO models are widely used to investigate the implication of a given gene in psychiatric disorders, they present several disadvantages [396]. First, in some cases inactivating genes that are essential to normal function leads to lethality during development. In this case, a possible solution is using conditional mutations that are specific of a given tissue or cell type or that are

limited in time, which was done in the case of *Rbfox1* KO mice. Second, full KO models are useful to study the function of a gene but do not represent exactly how variants in this gene with smaller effects may affect the phenotype. As mentioned before, the genetic basis of psychiatric disorders is usually formed by several variants with individual mild additive effects that act in combination. For this reason, it is also interesting to analyse HZ phenotypes, which mimic the effect of a partial downregulation of the gene. For this reason, *Bex3*, *Rbfox1* and *rbfox1* HET phenotypes were evaluated in our studies. Finally, in some cases KO of a single gene may not have any visible effect in the phenotype, or it may produce a phenotype different to the one expected. We should consider that in complex disorders the additive effect of variants in different genes is what gives rise to the phenotype. Moreover, in all our studies the genetic background has been reported as an important factor, similar to what occurs in humans, as variants in genes different from the one studied may contribute to the final phenotype. Further, compensation mechanisms may occur to balance the effects of the missing gene.

# 4. Inspecting brain activity and connectivity in zebrafish models

In the last years, the advances in fluorescence microscopy and the development of sensitive genetically encoded calcium indicators have converged to the development of whole-brain imaging, a recent non-invasive technique that permits obtaining *in vivo* recordings of the activity of single neurons in zebrafish larvae. Zebrafish larvae have the interesting advantage of being transparent and having a small size that allows exploration of the activity of single neurons in the whole brain *in vivo* [306,310]. In this thesis, we aimed to analyse the effect of *ywhaz* and *rbfox1* deficiency in neuronal activity and connectivity. To do so, we set up a whole-brain imaging protocol to analyse differences in neuronal activity between WT and mutant zebrafish larvae from our *ywhaz* and *rbfox1* KO lines. In the following sections I will discuss about the issues we experienced during the setup, as well as the results obtained.

## 4.1. Steps followed in the setup of the analysis

Creation of KO transgenic lines expressing GCaMP6s

To perform calcium brain imaging, larvae have to be completely transparent and express a calcium marker that allows detection of neuronal activity. An interesting advantage of calcium

indicators is that they enable to detect action potentials without the need of a high temporal resolution, as calcium transients have slower kinetics. In the last decades, a high number of genetically encoded calcium indicators (GECIs) have been developed with the interesting advantage of targeting specific cellular populations, the GCaMP family being the most used for *in vivo* recordings [397,398]. GCaMPs consist on a circularly permuted enhanced green fluorescent protein (cpGFP) linked to the calcium-binding protein calmodulin (CaM) and the CaM-binding peptide M13. The CaM/M13 complex is situated close to the chromophore inside cpGFP. In neurons, action potentials cause a rise in calcium concentration in the cytosol that produces conformational changes in the CaM/M13 complex and lead to an increase in fluorescence intensity (Figure 22A). This increase in fluorescence is reversed when calcium is subsequently pumped out from the cytosol [399]. Genetic engineering contributed to the optimization of the sensitivity and kinetics of the different GCaMPs markers and led to the development of the family of ultra-sensitive markers GCaMP6 [304,398].
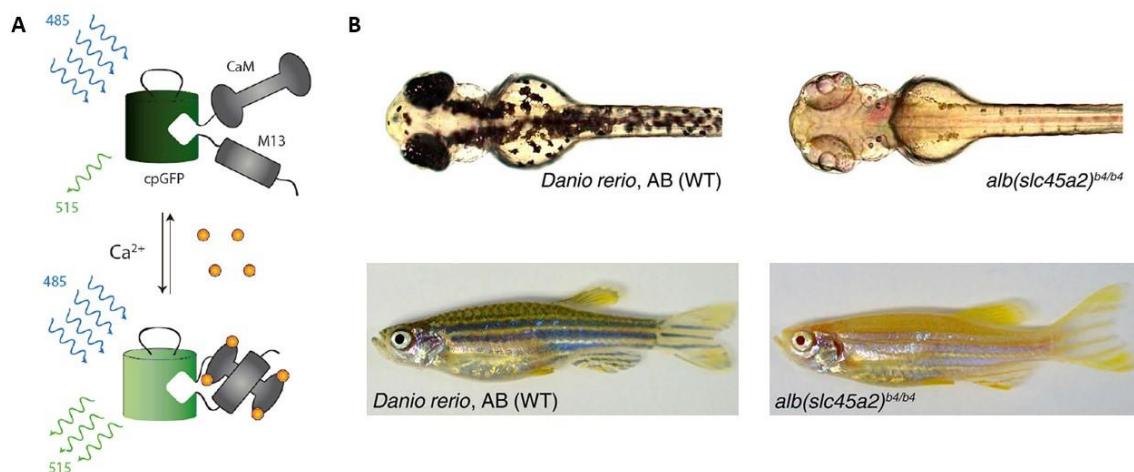


**Figure 22. GCaMP6s albino zebrafish line**. **A)** GCaMP markers are formed by a circularly permuted enhanced green fluorescent protein (cpGFP) linked to the calcium-binding protein calmodulin (CaM) and the CaM-binding peptide M13. Calcium-dependent conformational changes in the CaM/M13 complex lead to a reversible increase in cpGFP fluorescence intensity. **B)** On the left, AB wild-type (WT) zebrafish, that present pigmentation both in larval (3dpf) and adult stages. On the right, albino (*slc454a2* knockout) zebrafish that present no pigmentation during early larval stages (3dpf) and lack of melanin in adult stages. Adapted from Broussard et al., 2014 [399] and Antinucci and Hindges, 2016 [400].

In this work, to obtain transparent transgenic lines that are deficient in our candidate genes and express a calcium marker, we performed several crosses between albino fish expressing GCaMP6s pan-neuronally and KO fish for our candidate genes *ywhaz* and *rbfox1*. Compared to other GCaMP versions, GCaMP6s exhibits an excellent signal-to-noise ratio and presents a long half-decay rate that is ideal to measure action potentials at our frame rate, which is 1 Hz, as the whole brain is scanned once per second [398]. In addition, in zebrafish several cell populations

produce pigments that interfere with the optical accessibility to tissue while using microscopy. In this work, to reduce zebrafish skin pigmentation, we used a transgenic line deficient in the *slc45a2* gene, called albino line, in which melanin production is reduced and therefore the optical transparency is increased (Figure22B) [400,401].

<u>Set up of light-sheet microscopy recordings</u>

For the recordings, it was necessary to immobilise larvae to prevent movements. Tricaine (MS222) is used to immobilise zebrafish larvae in some contexts and present the advantage of a reduced economic cost. We used tricaine during the set up optimization, however, it was not useful for the final experiments as it is an anaesthetic that decreases neuronal activity. For the final experiments, we used α-bungarotoxine, a drug widely applied in imaging experiments with zebrafish larvae that produces paralysis by irreversibly blocking nicotinic acetylcholine receptors in the neuromuscular junction. Once we optimized the dose and time of paralysis with α-bungarotoxine, we adapted the light-sheet microscope objectives and refined the recording parameters – number of planes, frame rate, brain coverage and laser power. We also created a personalized plastic chamber with a plastic tube of a specific diameter in which larvae could be contained during the whole recording without moving. A limitation of our brain imaging recordings was the impossibility to record ventral brain areas with high resolution due to technical limitations of the set up.

<u>Movement correction</u>

Once the recordings of all the genotypes were obtained, we created a pipeline to correct the movement of the videos in which larvae presented a slight rotation throughout the recording. This slight rotation may be due to a smaller size of the brain in these fish, which may not fit properly in the plastic tube. We managed to correct almost all the recordings of WT and *ywhaz* fish, which allowed us to perform the subsequent analysis. However, the rotation was stronger in almost all *rbfox1* larvae, which may reflect differences in brain size probably related to an impaired brain development, and prevented us to continue with the analysis.

<u>Detection of neurons, extraction of fluorescence traces and deconvolution</u>

Several methods exist to detect neurons from calcium imaging recordings, some are based on automatic segmentation or morphology and others are based on fluorescence variations, the latter being more accurate. After unsuccessfully detect neurons with a pipeline based on neuronal morphology [402], we finally adapted a CaImAn pipeline [403] to detect single neurons and deconvolve their activity from the GCAMP6s indicator dynamics. This approach relies on a constrained nonnegative matrix factorization method (CNMF) [404] that expresses the

spatiotemporal fluorescence activity as the product of a spatial matrix, that encodes the location of each neuron, and a temporal matrix, that reflects the calcium concentration of each neuron over time. Several parameters should be considered to deconvolve neuronal activity from fluorescence traces, some that are related to the characteristics of the recordings, such as time and shape of fluorescence decay, bleaching, or size of neurons, and other related to computation capacity, such as lack of RAM memory. Once single-cell activity was extracted, we designed a MATLAB pipeline to manually apply a mask to each plane of the recording in order to avoid detecting artefacts outside the brain area.

Analysis of brain activity and connectivity

The last steps of the analysis involved Netcal (www.itsnetcal.com), a software we used to obtain neuronal activity and network dynamics statistics, and the Volume Segmenter, a MATLAB tool that was used to manually determine the different brains to be analysed. Importantly, Netcal allowed us to investigate not only single-cell activity, but also collective burst activity, a pattern that has a relevant role during early development. Also, several Advanced Normalization Tools (ANTs) exist to define brain areas in a recording based on a reference atlas, tools that are widely used in human MRI studies. However, we could not find ANTs that properly fit to our zebrafish data and we finally used the previously mentioned MATLAB tool and the Z-brain atlas as a reference (engertlab.fas.harvard.edu/Z-Brain/) to manually define the areas. Unfortunately, two of the five brain areas presented a low number of neurons in the majority of the individuals, which prevented us to perform the corresponding analyses there. Compared to other similar studies, our neuronal activity and connectivity analysis is more exhaustive, an aspect that will be discussed in the 4.4. section.

## 4.2.   The *ywhaz* model

The analysis of our whole-brain imaging recordings in WT and *ywhaz* KO larvae, created by CRISPR/Cas9 engineering, revealed an altered hindbrain spontaneous neuronal activity and functional connectivity in KO animals (Article 2). Specifically, our results point to a higher clustering and more effective connectivity in the cerebellum of KO fish, and to the presence of more isolated neuronal communities in the medulla oblongata of KO fish, which generates a lower collective burst activity and synchrony in this area. We hypothesize that these alterations in neuronal activity in specific areas might be due to an aberrant neuronal migration and synaptogenesis. Interestingly, *Ywhaz* KO mice present developmental abnormalities in specific brain areas, impaired neuronal migration and synaptic formation as well as reduced synaptic

density [220,222]. Also, early spontaneous synchronized burst activity is essential for the correct assembly of neural circuits during development, not only in mammals but also in zebrafish, as it influences neurogenesis, neuronal migration, synaptogenesis, apoptosis and myelination [65,405,406]. Importantly, spontaneous neuronal connectivity changes over development in zebrafish, showing a peak around 5-6 dpf [407], the age at which our analyses were performed. It has been demonstrated that an asynchronized activity during early development may disturb subsequent developmental mechanisms leading to miswiring and thus contribute to the onset of psychiatric conditions [65]. For this reason, we hypothesized that the reported alterations in hindbrain connectivity during larval stages may lead to long-term alterations in brain function.

## 4.3.   The *rbfox1* model

Unfortunately, we experienced issues when trying to correct movement in the recordings of the *rbfox1* line and, therefore, we need to readapt the pipeline to solve this problem. We hypothesize that *rbfox1* larvae have a smaller brain size, which would explain that larvae rotate more than WT and *ywhaz* larvae in the plastic tube during the recording. In mammals, *RBFOX1* is involved in neuronal migration and synapse network formation during corticogenesis [231], therefore, *rbfox1* deficiency in our zebrafish model may lead to impairments in neuronal migration during development and to differences in brain size. However, a proper analysis should be performed to confirm this hypothesis.

We also observed a high mortality among *rbfox1* mutants during the recordings, significantly more frequent than among WT and *ywhaz* KO fish. Interestingly, before brain death we observed a depolarization wave that lasted several minutes, started in specific lateral areas of the hindbrain and then spread all over the brain. *RBFOX1* deficiency has been related to an E/I imbalance, and especially to an increase in neuronal excitability [232,378]. Also, rare exonic deletions in *RBFOX1* have been described in patients with idiopathic generalized epilepsies (IGE) [408]. We therefore hypothesize that *rbfox1* deficit may cause an increased neuronal excitability in zebrafish larvae and lead to this seizure-like depolarization wave in some individuals.

## 4.4.   Benefits and limitations of whole-brain imaging

Calcium brain imaging is a non-invasive technique that emerged recently and allows *in vivo* analysis of broad patterns of activity with a single-cell resolution and without neuronal damage. Interestingly, this technique bridges the gap between fMRI and electrophysiological techniques.

In zebrafish, larvae transparency and small size allows recording of neural activity of the whole intact brain. In mammals, the scan of the whole brain is not possible but some techniques such as thinned-skull, cranial windows or microendoscope preparations allow to analyse neuronal activity in specific brain areas leaving the cells under study unperturbed [409,410]. This represents an important advantage compared to other electrophysiological methods that need electrode implantation and an invasive surgery. In addition, the use of GECI markers allows selective sampling of cellular subsets: all neurons in our case, but it is also possible to target subsets of neurons or other cellular types such as glia using different promoters [411].

Another advantage of this technique is the possibility to analyse single-cell activity and connectivity and observe thousands of neurons simultaneously *in vivo*. In humans, fMRI allows to detect differences in brain activity and functional connectivity, however, measures are based on oxygen levels, not in real neuronal activity, and do not permit the analysis of single cells but only of wide areas. Compared to fMRI, whole-brain imaging in animals allow a more refined analysis of brain activity and stands as a powerful tool to investigate neuronal correlates of psychiatric disorders in animal models.

In the last decade, whole-brain imaging has become a useful technique to investigate neuronal correlates of motor activity and response to stimuli in zebrafish but, only very recently, it has started to be used to analyse the effect of candidate genes in neuronal activity. To our knowledge, only two other studies to date have explored the effect of the deficiency of a given gene in neuronal circuitry [312,313]. Compared to these previous studies, we have performed a more comprehensive analysis of both neuronal single-cell and burst activity as well as connectivity within brain areas. In the study from Reichmann et al. [313], only 10 planes of the whole brain were scanned (we scanned 60) and single-cell activity was not analysed, only generalized fluorescence of defined brain areas, and therefore detailed information about spikes, firing, burst activity and detailed network connectivity was not available. In the study from Light et al. [312], although the analyses performed were more comprehensive (they analysed brain activity at different post-fertilization days), they only scanned 12 planes of the whole brain and therefore were only able to detect around 2,000 neurons, while we detected up to 15,000 neurons in some individuals, and did not analyse differences between brain areas. Nevertheless, a limitation of our study was the impossibility of scanning ventral areas of the brain due to a lack of resolution. It would have been interesting to analyse neuronal activity in deep ventral nuclei involved in dopaminergic and serotoninergic transmission and relevant in psychiatric disorders. In addition, we only analysed spontaneous brain activity, it would be

interesting to inspect neuronal circuitry of larvae exposed to external stimuli or in response to drugs, although changes in the setup would be needed to make it possible.

# 5. Gaining insight into the genetic architecture of neurodevelopmental and psychiatric disorders

All the analyses and experiments performed in this thesis have contributed to a better understanding of the function of candidate genes and their implication in neurodevelopmental and psychiatric disorders. Below, I will briefly summarize our main findings and how they contribute to gaining insight into the genetic architecture of mental disorders.

## 5.1. The 14-3-3 gene family

Our results highlight the contribution of common variants in the 14-3-3 gene family, especially in *YWHAE,* to SCZ. On the other hand, we observe that rare variants in 14-3-3 seem to have a higher contribution to ASD, as we find a significant burden of URVs in ASD patients compared to controls, being URVs three times more frequent in patients. In addition, damaging rare variants in *YWHAZ* may contribute to both ASD and SCZ. First, a significant burden of rare SNVs in *YWHAZ* was identified in SCZ patients compared to controls. Also, we found that a truncating mutation in *YWHAZ* leads to protein loss of function, which probably contributes to the ASD and ADHD phenotypes described in the affected siblings. This hypothesis was further supported by our results in a zebrafish KO model, in which *ywhaz* deficiency affects brain connectivity during development and produces neurotransmitter and behavioural alterations similar to those described in ASD and SCZ patients.

## *5.2. RBFOX1*

We report an important contribution of common variants in *RBFOX1* to several psychiatric disorders, suggesting its pleiotropic effects in psychiatry. Interestingly, the rs6500744 variant, previously related to aggression, was found to alter brain circuitry involved in integrating cognition with emotion, executive function and impulsivity. All these results suggest that common variation in *RBFOX1* may contribute to an increased risk of psychiatric disorders. On the other hand, we have reported a high number of CNVs spanning *RBFOX1* non-coding regions,

probably regulatory regions, in patients with different psychiatric disorders, especially ASD and SCZ. Furthermore, our *RBFOX1* KO animal models present strong behavioural alterations that resemble symptoms described in ASD and ADHD patients, reinforcing the contribution of damaging rare variation in *RBFOX1* to neurodevelopmental disorders.

## 5.3. *BEX/TCEAL*

We could not report any common variant in *BEX3* significantly associated with ASD or SCZ, although the poor quality of X chromosome in GWAS data may have impacted our results. Also, we did not identify rare variants spanning *BEX3* specifically, but rare CNVs spanning several *BEX/TCEAL* genes had been previously reported in patients with severe neurodevelopmental problems. Our mice models of *Bex3* deficiency present morphological and electrophysiological alterations that, together with the behavioural changes described, point to a contribution of *Bex3* deficiency, modulated by other genetic factors, to neurodevelopmental disorders.

## 5.4. The genetic architecture of neurodevelopmental and psychiatric disorders

The first genetic studies in ASD focused mainly on rare variants, whereas in other psychiatric disorders such as ADHD or SCZ in the initial focus was on common variation. Later, the improvement of genetic analyses highlighted the complexity of genetic factors underlying these disorders and the polygenic contribution of both rare and common variants to their aetiology. Indeed, recent genome-wide analyses have helped to identify hundreds of genes and variants contributing to psychiatric and neurodevelopmental disorders, pointing to an omnigenic model in psychiatry [412].

The results obtained in this thesis in regard of the candidate genes studied point to a differential contribution of common and rare variants in these genes to the outcoming phenotypes. Common variation in these genes seems to have a wider effect in psychiatric disorders, whereas rare variation seems to lead to more severe phenotypes related to impaired neurodevelopment. Importantly, using different animal models we have demonstrated that the genetic background has an influence in the downstream effects of the loss of function of a given gene, resulting in similar phenotypes but with slight differences. In humans, these slight differences may lead to a different or multiple diagnosis because, as mentioned previously, there is a high overlap in symptoms across psychiatric and neurodevelopmental disorders, and sometimes the clinical borders between them are unclear (Figure 23). Taken together, these results reinforce a recent

view in psychiatric genetics that suggests a major role of pleiotropy in mental disorders and challenges the current diagnostic boundaries defined by the DSM manual [413].
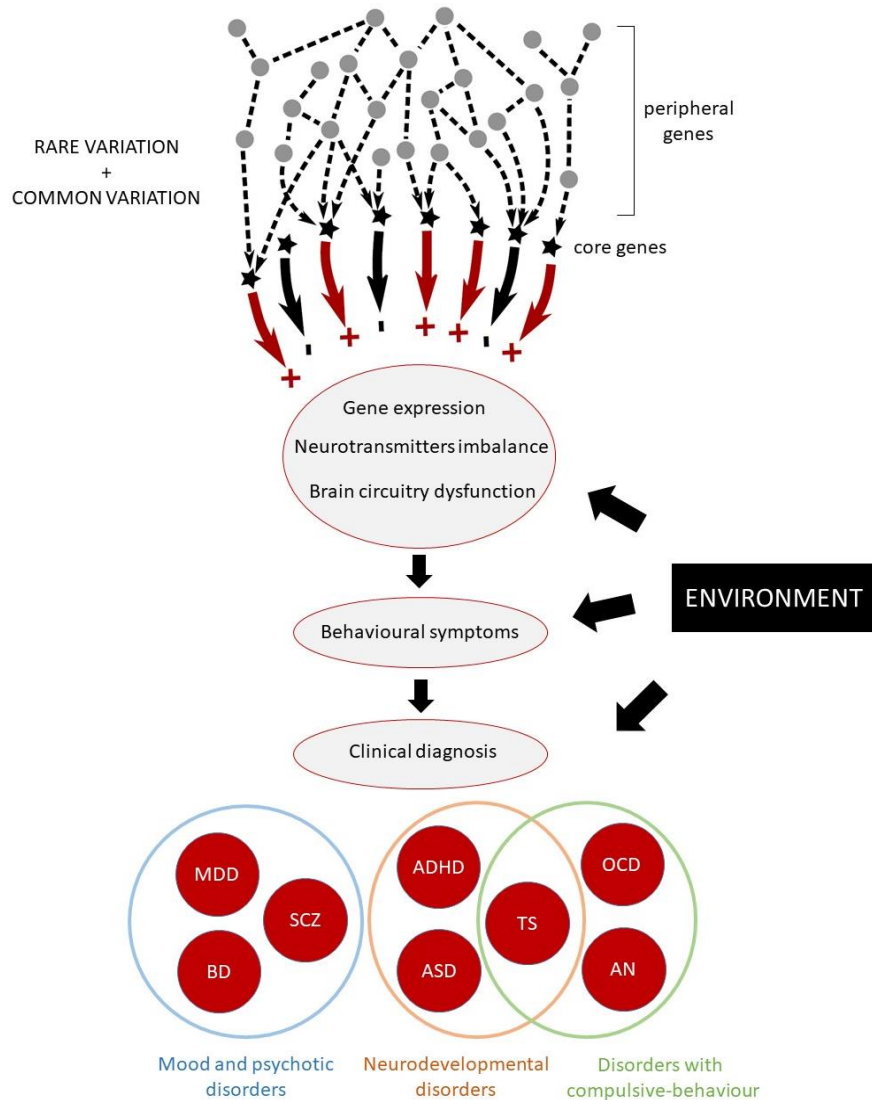


**Figure 23. Proposed omnigenic model of complex psychiatric and neurodevelopmental disorders.** Rare variation in peripheral and core genes contribute to alterations in gene expression, neurotransmitter systems and brain circuitry. These alterations lead to behavioural symptoms that determine the clinical diagnosis following predefined symptomatic criteria. Here, the eight disorders investigated in the last cross-disorder meta-analysis by Lee are represented, although other mental disorders are also concerned in this model. The categories in which the disorders are included are based on the genetic correlations observed by Lee et al. [192]. ADHD, attention deficit/hyperactivity disorder; AN, anorexia; ASD, autism spectrum disorder; BD, bipolar disorder; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; SCZ, schizophrenia; TS, Tourette Syndrome. Adapted from Liu et al., 2019 and Gandal et al., 2018 [188,414].

In psychiatry, the co-occurrence of several conditions in the same individual seems to be the rule rather than the exception. Different mechanism of pleiotropy have been described to explain psychiatric comorbidities [412]. First, vertical or biological pleiotropy occurs when shared genetic causes explain the comorbidity (one or several genes can be involved). In this case, variations in the same gene lead to a wide range of phenotypes depending on the effect size of the variants and on the additive contribution of variants present in other genes (Figure 24A). Second, horizontal or mediated pleiotropy occurs when a secondary disorder appears as a consequence of a feature of a primary disorder, e.g. substance use disorders may be triggered by one of the traits that are characteristic of ADHD, such as impulsivity (Figure 24B). Finally, apparent pleiotropic associations may arise from genetic studies although different genetic mechanisms underlie the two disorders (spurious pleiotropy) (Figure 24C).
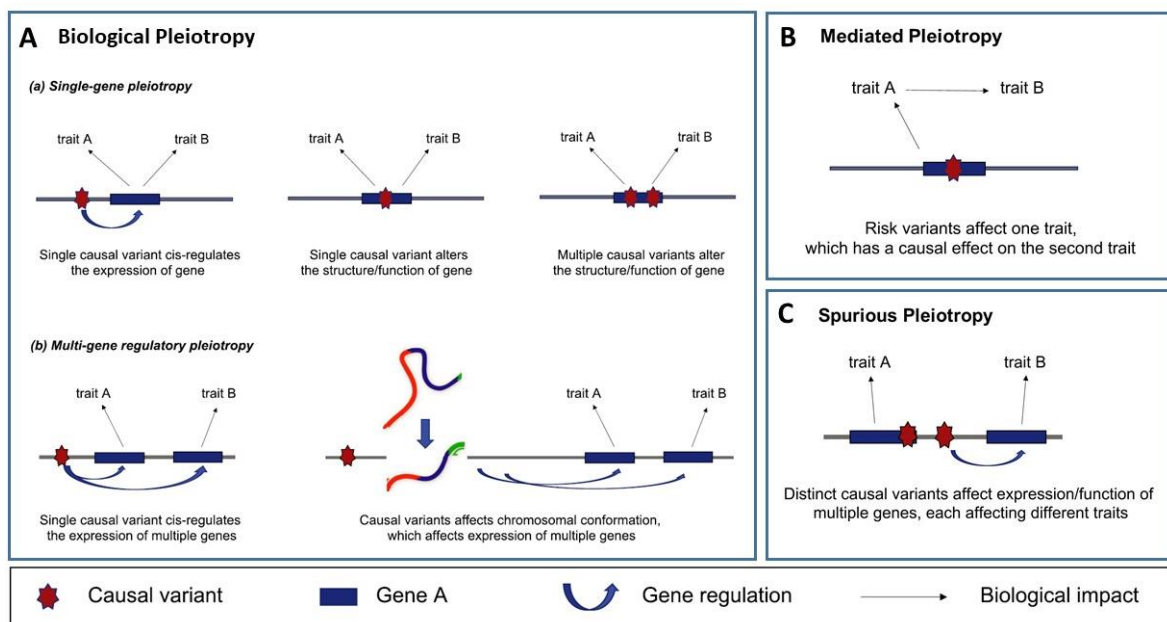


**Figure 24. Pleiotropic mechanism underlying cross-phenotype associations**. **A)** Biological pleiotropy includes (a) single-gene pleiotropy, in which causal variants affect the function, activity or expression of a single gene that underlies one or more traits; and (b) multi-gene regulatory pleiotropy, in which noncoding causal variants affect the expression of multiple genes that may underlie different traits. **B)** In mediated pleiotropy, causal variants influence one trait that causes a second trait. **C)** Spurious pleiotropy refers to situations when cross-trait associations are described due to limitations in the study design although the underlying genetic causes are different. Adapted from Lee et al., 2021 [412].

In line with biological pleiotropic mechanisms contributing to psychiatry, two recent meta-analyses found strong genetic correlations among psychiatric disorders in terms of common

variation [192,193] (Figure 25). In the work from Lee et al., ASD showed a significant genetic correlation with ADHD, BD, MDD and SCZ [192] (Figure 25B). We are still far from a complete understanding of psychiatric genetics. However, progress in genome-wide analyses with the increasing availability of large-scale genomic and phenotypic databases, together with technological improvements in cellular and animal models, will help in a near future to gain insight into the complex genetic factors that underlie psychiatric disorders. Importantly, a better knowledge of the biological and molecular mechanisms behind psychiatric diseases would facilitate a better diagnosis and treatment.
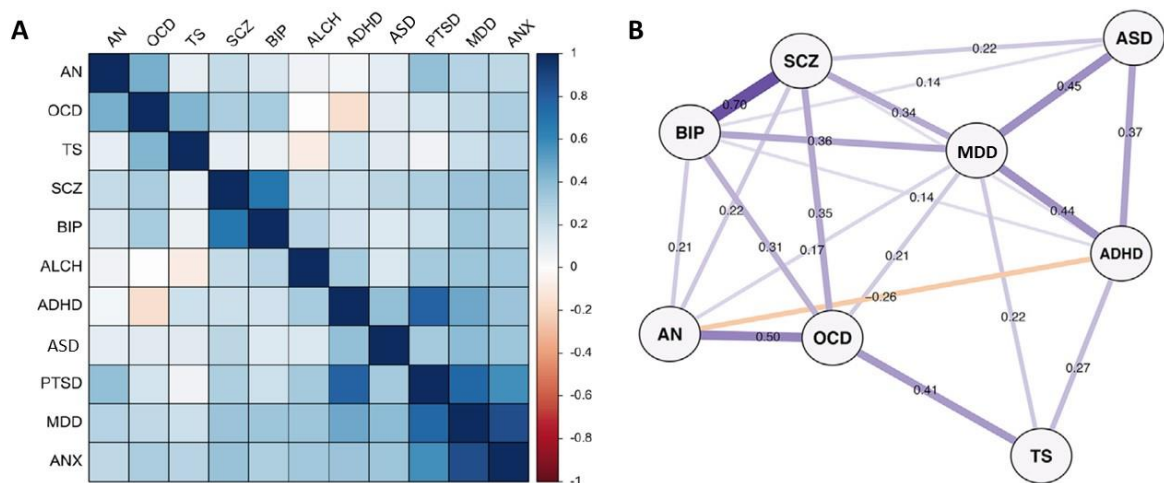


F**igure 25. Genetic correlations found between psychiatric disorders. A)** Heatmap of genetic correlations estimated between 11 psychiatric disorders using linkage disequilibrium score regression (LDSC) in the Grotzinger et al. meta-analysis [193]. **B)** SNP-based genetic correlations overcoming Bonferroni correction in the Lee et al. meta-analysis [192]. Each node represents a disorder and edges indicate the strength of the correlations. ADHD, attention deficit/hyperactivity disorder; ALCH, problematic alcohol use; AN, anorexia; ANX, anxiety; ASD, autism spectrum disorder; BIP, bipolar disorder; MDD, major depressive disorder; OCD, obsessive-compulsive disorder; PTSD, post-traumatic stress disorder; SCZ, schizophrenia; TS, Tourette Syndrome. Adapted from Grotzinger et al., [193] and Lee et al., [192].

# CONCLUSIONS

14-3-3, *RBFOX1* and *BEX*/*TCEAL* genes exert pleiotropic effects on ASD and comorbid psychiatric disorders by impacting neurodevelopment, as shown by genetic studies in humans and animal models.

**Exploring the contribution of the 14-3-3 gene family to ASD and other psychiatric disorders**

1. Common variants in the 14-3-3 genes contribute to schizophrenia (SCZ) whereas rare variants in these genes are enriched in ASD and SCZ.

2. The 14-3-3 genes show altered expression in *postmortem* brain areas of ASD or schizophrenia patients.

3. The p.L220Ffs*18 *YWHAZ* rare variant, present in two siblings with ASD and ADHD, has a loss of function effect, disrupting protein dimerization and binding to tyrosine hydroxylase.

4. In zebrafish, *ywhaz* shows a pan-neuronal expression during development and a restricted expression in adult Purkinje cells, a region affected in ASD patients.

5. *ywhaz* depletion causes an altered activity and connectivity in the hindbrain during development, and an altered monoamine signalling in the hindbrain in zebrafish adults that leads to a freezing behaviour that can be rescued with specific drug treatments.

**Exploring the contribution of *RBFOX1* to ASD and other psychiatric disorders**

6. Common variants in *RBFOX1* are associated with eight psychiatric traits and CNVs spanning *RBFOX1* are reported in patients with different psychiatric conditions, particularly in ASD or schizophrenia.

7. *RBFOX1* shows a significant decreased expression in *postmortem* cortex of ASD and schizophrenia patients.

8. The rs6500744 variant situated in *RBFOX1* alters brain circuitry involved in integrating cognition with emotion, executive function and impulsivity.

9. *Rbfox1*-deficient mice present behavioural alterations that can be related to neurodevelopmental disorders such as ASD, ADHD and SCZ.

10. In zebrafish, *rbfox1* is pan-neuronally expressed during development and shows a restricted expression in adult forebrain areas, regions that play a role in sensory processing, stress and behaviour.

11. Both *rbfox1*-deficient zebrafish lines present hyperactivity, thigmotaxis, and alterations in social interaction, considered ASD- and ADHD-like behaviours.

**Exploring the contribution of *BEX/TCEAL* gene family to ASD and other psychiatric disorders**

12. All *BEX/TCEAL* genes are downregulated in *postmortem* brain regions of ASD and schizophrenia patients

13. *Bex3*-deficient mice show anatomical and molecular brain alterations, an excitatory/inhibitory imbalance, and ASD- and SCZ-like behaviours.

# REFERENCES

1. Bleuler, P.E. *Dementia Praecox, or the Group of Schizophrenias*; **1911**;

2. Kanner, L. Autistic disturbances of affective contact. *Nerv. Child* **1943**, *2*, 217–250.

3. Asperger, H. Die "Autistischen Psychopathen" im Kindesalter. *Arch. Psychiatr. Nervenkr.* **1944**, *117*, 76–136, doi:doi:10.1007/BF01837709.

4. American Psychiatric Association *DSM-I: Mental disorders; diagnostic and statistical manual*; **1952**;

5. American Psychiatric Association DSM-II: Diagnostic and Statistical Manual of Mental Disorders, 2nd Edition. **1968**.

6. Bettelheim, B. *The empty fortress : infantile autism and the birth of the self*; **1967**; ISBN 0029031400.

7. American Psychiatric Association *DSM-III: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition*; **1980**;

8. American Psychiatric Association *DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*; **1994**;

9. American Psychiatric Association *DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*; **2000**;

10. American Psychiatric Association *DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*; **2013**;

11. Baxter, A.J.; Brugha, T.S.; Erskine, H.E.; Scheurer, R.W.; Vos, T.; Scott, J.G. The epidemiology and global burden of autism spectrum disorders. *Psychol. Med.* **2015**, *45*, 601–613, doi:10.1017/S003329171400172X.

12. Elsabbagh, M.; Divan, G.; Koh, Y.J.; Kim, Y.S.; Kauchali, S.; Marcín, C.; Montiel-Nava, C.; Patel, V.; Paula, C.S.; Wang, C.; et al. Global Prevalence of Autism and Other Pervasive Developmental Disorders. *Autism Res.* **2012**, *5*, 160–179, doi:10.1002/aur.239.

13. Singh, J.S.; Bunyak, G. Autism Disparities: A Systematic Review and Meta-Ethnography of Qualitative Research. *Qual. Health Res.* **2019**, *29*, 796–808, doi:10.1177/1049732318808245.

14. Brugha, T.S.; Spiers, N.; Bankart, J.; Cooper, S.A.; McManus, S.; Scott, F.J.; Smith, J.; Tyrer, F. Epidemiology of autism in adults across age groups and ability levels. *Br. J. Psychiatry* **2016**, *209*, 498–503, doi:10.1192/bjp.bp.115.174649.

15. Loomes, R.; Hull, L.; Mandy, W.P.L. What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **2017**, *56*, 466–474.

16. Ferri, S.L.; Abel, T.; Brodkin, E.S. Sex Differences in Autism Spectrum Disorder: a Review. *Curr. Psychiatry Rep.* **2018**, *20*, doi:10.1007/s11920-018-0874-2.

17. Russell, G.; Steer, C.; Golding, J. Social and demographic factors that influence the diagnosis of autistic spectrum disorders. *Soc. Psychiatry Psychiatr. Epidemiol.* **2011**, *46*, 1283–1293, doi:10.1007/s00127-010-0294-z.

18. Constantino, J.N.; Charman, T. Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. *Lancet Neurol.* **2016**, *15*, 279–291, doi:10.1016/S1474-4422(15)00151-9.

19. Dworzynski, K.; Ronald, A.; Bolton, P.; Happé, F. How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *J. Am. Acad. Child Adolesc. Psychiatry* **2012**, *51*, 788–797, doi:10.1016/j.jaac.2012.05.018.

20. Begeer, S.; Mandell, D.; Wijnker-Holmes, B.; Venderbosch, S.; Rem, D.; Stekelenburg, F.; Koot, H.M. Sex

differences in the timing of identification among children and adults with autism spectrum disorders. *J. Autism Dev. Disord.* **2013**, *43*, 1151–1156, doi:10.1007/s10803-012-1656-z.

21. Lai, M.C.; Lombardo, M. V.; Baron-Cohen, S. Autism. *Lancet* **2014**, *383*, 896–910, doi:10.1016/S0140-6736(13)61539-1.

22. Lord, C.; Brugha, T.S.; Charman, T.; Cusack, J.; Dumas, G.; Frazier, T.; Jones, E.J.H.; Jones, R.M.; Pickles, A.; State, M.W.; et al. Autism spectrum disorder. *Nat. Rev. Dis. Prim.* **2020**, *6*, 5, doi:10.1038/s41572-019-0138-4.

23. Lord, C.; Rutter, M.; Le Couteur, A. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **1994**, *24*, 659–685, doi:10.1007/BF02172145.

24. Lord, C.; Rutter, M.; DiLavore, P.C.; Risi, S.; Gotham, K.; Bishop, S.L.; Luyster, R.J.; Guthrie, W. *ADOS-2: Autism diagnostic observation schedule, Second Edition*; **2012**;

25. Charman, T.; Gotham, K. Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practise. *Child Adolesc. Ment. Health* **2013**, *18*, 52–63, doi:10.1111/j.1475-3588.2012.00664.x.

26. Chlebowski, C.; Robins, D.L.; Barton, M.L.; Fein, D. Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics* **2013**, *131*, doi:10.1542/peds.2012-1525.

27. Stenberg, N.; Bresnahan, M.; Gunnes, N.; Hirtz, D.; Hornig, M.; Lie, K.K.; Lipkin, W.I.; Lord, C.; Magnus, P.; Reichborn-Kjennerud, T.; et al. Identifying children with autism spectrum disorder at 18 months in a general population sample. *Paediatr. Perinat. Epidemiol.* **2014**, *28*, 255–262, doi:10.1111/ppe.12114.

28. Kleinman, J.M.; Robins, D.L.; Ventola, P.E.; Pandey, J.; Boorstein, H.C.; Esser, E.L.; Wilson, L.B.; Rosenthal, M.A.; Sutera, S.; Verbalis, A.D.; et al. The modified checklist for autism in toddlers: A follow-up study investigating the early detection of autism spectrum disorders. *J. Autism Dev. Disord.* **2008**, *38*, 827–839, doi:10.1007/s10803-007-0450-9.

29. Rosen, T.E.; Mazefsky, C.A.; Vasa, R.A.; Lerner, M.D. Co-occurring psychiatric conditions in autism spectrum disorder. *Int. Rev. Psychiatry* **2018**, *30*, 40–61, doi:10.1080/09540261.2018.1450229.

30. Canitano, R. Epilepsy in autism spectrum disorders. *Eur. Child Adolesc. Psychiatry* **2007**, *16*, 61–66, doi:10.1007/s00787-006-0563-2.

31. Besag, F.M.C.; Vasey, M.J. Seizures and Epilepsy in Autism Spectrum Disorder. *Child Adolesc. Psychiatr. Clin. N. Am.* **2020**, *29*, 483–500, doi:10.1016/j.chc.2020.02.002.

32. Yasuhara, A. Correlation between EEG abnormalities and symptoms of autism spectrum disorder (ASD). *Brain Dev.* **2010**, *32*, 791–798, doi:10.1016/j.braindev.2010.08.010.

33. Rommelse, N.N.J.; Geurts, H.M.; Franke, B.; Buitelaar, J.K.; Hartman, C.A. A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neurosci. Biobehav. Rev.* **2011**, *35*, 1363–1396, doi:10.1016/j.neubiorev.2011.02.015.

34. Craig, F.; Margari, F.; Legrottaglie, A.R.; Palumbi, R.; de Giambattista, C.; Margari, L. A review of executive function deficits in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Neuropsychiatr. Dis. Treat.* **2016**, *12*, 1191–1202, doi:10.2147/NDT.S104620.

35. Antshel, K.M.; Zhang-James, Y.; Wagner, K.E.; Ledesma, A.; Faraone, S. V. An update on the comorbidity of ADHD

and ASD: A focus on clinical management. *Expert Rev. Neurother.* **2016**, *16*, 279–293, doi:10.1586/14737175.2016.1146591.

36. Chisholm, K.; Lin, A.; Abu-Akel, A.; Wood, S.J. The association between autism and schizophrenia spectrum disorders: A review of eight alternate models of co-occurrence. *Neurosci. Biobehav. Rev.* **2015**, *55*, 173–183, doi:10.1016/j.neubiorev.2015.04.012.

37. Hill, A.P.; Zuckerman, K.E.; Hagen, A.D.; Kriz, D.J.; Duvall, S.W.; Van Santen, J.; Nigg, J.; Fair, D.; Fombonne, E. Aggressive behavior problems in children with autism spectrum disorders: Prevalence and correlates in a large clinical sample. *Res. Autism Spectr. Disord.* **2014**, *8*, 1121–1133, doi:10.1016/j.rasd.2014.05.006.

38. Hazlett, H.C.; Gu, H.; Munsell, B.C.; Kim, S.H.; Styner, M.; Wolff, J.J.; Elison, J.T.; Swanson, M.R.; Zhu, H.; Botteron, K.N.; et al. Early brain development in infants at high risk for autism spectrum disorder. *Nature* **2017**, *542*, 348–351, doi:10.1038/nature21369.

39. Courchesne, E.; Pierce, K.; Schumann, C.M.; Redcay, E.; Buckwalter, J.A.; Kennedy, D.P.; Morgan, J. Mapping early brain development in autism. *Neuron* 2007, *56*, 399–413.

40. Sacco, R.; Gabriele, S.; Persico, A.M. Head circumference and brain size in autism spectrum disorder: A systematic review and meta-analysis. *Psychiatry Res. - Neuroimaging* **2015**, *234*, 239–251, doi:10.1016/j.pscychresns.2015.08.016.

41. Carper, R.A.; Courchesne, E. Localized enlargement of the frontal cortex in early autism. *Biol. Psychiatry* **2005**, *57*, 126–133, doi:10.1016/j.biopsych.2004.11.005.

42. Stanfield, A.C.; McIntosh, A.M.; Spencer, M.D.; Philip, R.; Gaur, S.; Lawrie, S.M. Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *Eur. Psychiatry* 2008, *23*, 289–299.

43. Rojas, D.C.; Peterson, E.; Winterrowd, E.; Reite, M.L.; Rogers, S.J.; Tregellas, J.R. Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. *BMC Psychiatry* **2006**, *6*, doi:10.1186/1471-244X-6-56.

44. Schumann, C.M.; Hamstra, J.; Goodlin-Jones, B.L.; Lotspeich, L.J.; Kwon, H.; Buonocore, M.H.; Lammers, C.R.; Reiss, A.L.; Amaral, D.G. The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *J. Neurosci.* **2004**, *24*, 6392–6401, doi:10.1523/JNEUROSCI.1297-04.2004.

45. Stanfield, A.C.; McIntosh, A.M.; Spencer, M.D.; Philip, R.; Gaur, S.; Lawrie, S.M. Towards a neuroanatomy of autism: A systematic review and meta-analysis of structural magnetic resonance imaging studies. *Eur. Psychiatry* 2008, *23*, 289–299.

46. Rojas, D.C.; Peterson, E.; Winterrowd, E.; Reite, M.L.; Rogers, S.J.; Tregellas, J.R. Regional gray matter volumetric changes in autism associated with social and repetitive behavior symptoms. *BMC Psychiatry* **2006**, *6*, doi:10.1186/1471-244X-6-56.

47. Schumann, C.M.; Hamstra, J.; Goodlin-Jones, B.L.; Lotspeich, L.J.; Kwon, H.; Buonocore, M.H.; Lammers, C.R.; Reiss, A.L.; Amaral, D.G. The amygdala is enlarged in children but not adolescents with autism; the hippocampus is enlarged at all ages. *J. Neurosci.* **2004**, *24*, 6392–6401, doi:10.1523/JNEUROSCI.1297-04.2004.

48. Just, M.A.; Cherkassky, V.L.; Keller, T.A.; Kana, R.K.; Minshew, N.J. Functional and anatomical cortical underconnectivity in autism: Evidence from an fmri study of an executive function task and corpus callosum morphometry. *Cereb. Cortex* **2007**, *17*, 951–961, doi:10.1093/cercor/bhl006.

49. Courchesne, E.; Mouton, P.R.; Calhoun, M.E.; Semendeferi, K.; Ahrens-Barbeau, C.; Hallet, M.J.; Barnes, C.C.; Pierce, K. Neuron number and size in prefrontal cortex of children with autism. *JAMA - J. Am. Med. Assoc.* **2011**, *306*, 2001–2010, doi:10.1001/jama.2011.1638.

50. Courchesne, E.; Pramparo, T.; Gazestani, V.H.; Lombardo, M. V.; Pierce, K.; Lewis, N.E. The ASD Living Biology: from cell proliferation to clinical phenotype. *Mol. Psychiatry* **2019**, *24*, 88–107, doi:10.1038/s41380-018-0056-y.

51. Courchesne, E.; Karns, C.M.; Davis, H.R.; Ziccardi, R.; Carper, R.A.; Tigue, Z.D.; Chisum, H.J.; Moses, P.; Pierce, K.; Lord, C.; et al. Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology* **2001**, *57*, 245–254, doi:10.1212/WNL.57.2.245.

52. Hallahan, B.; Daly, E.M.; McAlonan, G.; Loth, E.; Toal, F.; O'Brien, F.; Robertson, D.; Hales, S.; Murphy, C.; Murphy, K.C.; et al. Brain morphometry volume in autistic spectrum disorder: A magnetic resonance imaging study of adults. *Psychol. Med.* **2009**, *39*, 337–346, doi:10.1017/S0033291708003383.

53. Scott, J.A.; Schumann, C.M.; Goodlin-Jones, B.L.; Amaral, D.G. A comprehensive volumetric analysis of the cerebellum in children and adolescents with autism spectrum disorder. *Autism Res.* **2009**, *2*, 246–257, doi:10.1002/aur.97.

54. Courchesne, E.; Yeung-Courchesne, R.; Press, G.A.; Hesselink, J.R.; Jernigan, T.L. Hypoplasia of cerebellar vermal lobules VI and VII in autism. In *The Science of Mental Health: Volume 2: Autism*; Taylor and Francis, 2013; Vol. 318, pp. 197–202 ISBN 9781136800818.

55. Courchesne, E.; Karns, C.M.; Davis, H.R.; Ziccardi, R.; Carper, R.A.; Tigue, Z.D.; Chisum, H.J.; Moses, P.; Pierce, K.; Lord, C.; et al. Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology* **2001**, *57*, 245–254, doi:10.1212/WNL.57.2.245.

56. Stoodley, C.J. Distinct regions of the cerebellum show gray matter decreases in autism, ADHD, and developmental dyslexia. *Front. Syst. Neurosci.* **2014**, *8*, doi:10.3389/fnsys.2014.00092.

57. D'Mello, A.M.; Crocetti, D.; Mostofsky, S.H.; Stoodley, C.J. Cerebellar gray matter and lobular volumes correlate with core autism symptoms. *NeuroImage Clin.* **2015**, *7*, 631–639, doi:10.1016/j.nicl.2015.02.007.

58. Fatemi, S.H.; Halt, A.R.; Realmuto, G.; Earle, J.; Kist, D.A.; Thuras, P.; Metz, A. Purkinje cell size is reduced in cerebellum of patients with autism. *Cell. Mol. Neurobiol.* **2002**, *22*, 171–175, doi:10.1023/A:1019861721160.

59. Whitney, E.R.; Kemper, T.L.; Rosene, D.L.; Bauman, M.L.; Blatt, G.J. Density of cerebellar basket and stellate cells in autism: Evidence for a late developmental loss of Purkinje cells. *J. Neurosci. Res.* **2009**, *87*, 2245–2254, doi:10.1002/jnr.22056.

60. Courchesne, E.; Pierce, K. Brain overgrowth in autism during a critical time in development: Implications for frontal pyramidal neuron and interneuron development and connectivity. *Int. J. Dev. Neurosci.* **2005**, *23*, 153–170, doi:10.1016/j.ijdevneu.2005.01.003.

61. Wegiel, J.; Kuchna, I.; Nowicki, K.; Imaki, H.; Wegiel, J.; Marchi, E.; Ma, S.Y.; Chauhan, A.; Chauhan, V.; Bobrowicz, T.W.; et al. The neuropathology of autism: Defects of neurogenesis and neuronal migration, and dysplastic changes. *Acta Neuropathol.* **2010**, *119*, 755–770, doi:10.1007/s00401-010-0655-4.

62. Lammert, D.B.; Howell, B.W. RELN mutations in autism spectrum disorder. *Front. Cell. Neurosci.* **2016**, *10*, 1–9, doi:10.3389/fncel.2016.00084.

63. Reiner, O.; Karzbrun, E.; Kshirsagar, A.; Kaibuchi, K. Regulation of neuronal migration, an emerging topic in

autism spectrum disorders. *J. Neurochem.* **2016**, *136*, 440–456, doi:10.1111/jnc.13403.

64. Martínez-Cerdeño, V. Dendrite and spine modifications in autism and related neurodevelopmental disorders in patients and animal models. *Dev. Neurobiol.* **2017**, *77*, 393–404, doi:10.1002/dneu.22417.

65. Molnár, Z.; Luhmann, H.J.; Kanold, P.O. Transient cortical circuits match spontaneous and sensory-driven activity during development. *Science (80-. ).* **2020**, *370*, doi:10.1126/science.abb2153.

66. Leighton, A.H.; Lohmann, C. The wiring of developing sensory circuits—From patterned spontaneous activity to synaptic plasticity mechanisms. *Front. Neural Circuits* **2016**, *10*, 1–13, doi:10.3389/fncir.2016.00071.

67. Toga, A.W.; Thompson, P.M. Mapping brain asymmetry. *Nat. Rev. Neurosci.* **2003**, *4*, 37–48, doi:10.1038/nrn1009.

68. De Schotten, M.T.; Dell'Acqua, F.; Forkel, S.J.; Simmons, A.; Vergani, F.; Murphy, D.G.M.; Catani, M. A lateralized brain network for visuospatial attention. *Nat. Neurosci.* **2011**, *14*, 1245–1246, doi:10.1038/nn.2905.

69. Lindell, A. *Lateralization of the expression of facial emotion in humans*; 1st ed.; Elsevier B.V., 2018; Vol. 238; ISBN 9780128146712.

70. Hodgson, J.C.; Hudson, J.M. *Speech lateralization and motor control*; 1st ed.; Elsevier B.V., 2018; Vol. 238; ISBN 9780128146712.

71. Eyler, L.T.; Pierce, K.; Courchesne, E. A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism. *Brain* **2012**, *135*, 949–960, doi:10.1093/brain/awr364.

72. Nielsen, J.A.; Zielinski, B.A.; Fletcher, P.T.; Alexander, A.L.; Lange, N.; Bigler, E.D.; Lainhart, J.E.; Anderson, J.S. Abnormal lateralization of functional connectivity between language and default mode regions in autism. *Mol. Autism* **2014**, *5*, 1–11, doi:10.1186/2040-2392-5-8.

73. Postema, M.C.; van Rooij, D.; Anagnostou, E.; Arango, C.; Auzias, G.; Behrmann, M.; Filho, G.B.; Calderoni, S.; Calvo, R.; Daly, E.; et al. Altered structural brain asymmetry in autism spectrum disorder in a study of 54 datasets. *Nat. Commun.* **2019**, *10*, 1–12, doi:10.1038/s41467-019-13005-8.

74. Opris, I.; Casanova, M.F. Prefrontal cortical minicolumn: From executive control to disrupted cognitive processing. *Brain* 2014, *137*, 1863–1875.

75. Casanova, M.F.; van Kooten, I.A.J.; Switala, A.E.; van Engeland, H.; Heinsen, H.; Steinbusch, H.W.M.; Hof, P.R.; Trippe, J.; Stone, J.; Schmitz, C. Minicolumnar abnormalities in autism. *Acta Neuropathol.* **2006**, *112*, 287–303, doi:10.1007/s00401-006-0085-5.

76. Casanova, M.F.; Buxhoeveden, D.; Gomez, J. Disruption in the Inhibitory Architecture of the Cell Minicolumn: Implications for Autisim. *Neuroscientist* 2003, *9*, 496–507.

77. O'Reilly, C.; Lewis, J.D.; Elsabbagh, M. Is functional brain connectivity atypical in autism? A systematic review of EEG and MEG studies. *PLoS One* **2017**, *12*, 1–28, doi:10.1371/journal.pone.0175870.

78. Rane, P.; Cochran, D.; Hodge, S.M.; Haselgrove, C.; Kennedy, D.N.; Frazier, J.A. Connectivity in Autism: A Review of MRI Connectivity Studies. *Harv. Rev. Psychiatry* 2015, *23*, 223–244.

79. Khan, A.J.; Nair, A.; Keown, C.L.; Datko, M.C.; Lincoln, A.J.; Müller, R.A. Cerebro-cerebellar resting-state functional connectivity in children and adolescents with autism spectrum disorder. *Biol. Psychiatry* **2015**, *78*, 625–634, doi:10.1016/j.biopsych.2015.03.024.

80. Su, L. Da; Xu, F.X.; Wang, X.T.; Cai, X.Y.; Shen, Y. Cerebellar Dysfunction, Cerebro-cerebellar Connectivity and

Autism Spectrum Disorders. *Neuroscience* **2021**, *462*, 320–327, doi:10.1016/j.neuroscience.2020.05.028.

81.  Atladóttir, H.Ó.; Pedersen, M.G.; Thorsen, P.; Mortensen, P.B.; Deleuran, B.; Eaton, W.W.; Parner, E.T. Association of family history of autoimmune diseases and autism spectrum disorders. *Pediatrics* **2009**, *124*, 687–694, doi:10.1542/peds.2008-2445.

82.  Siniscalco, D.; Schultz, S.; Brigida, A.L.; Antonucci, N. Inflammation and neuro-immune dysregulations in autism spectrum disorders. *Pharmaceuticals* **2018**, *11*, 1–14, doi:10.3390/ph11020056.

83.  Onore, C.; Careaga, M.; Ashwood, P. The role of immune dysfunction in the pathophysiology of autism. *Brain. Behav. Immun.* **2012**, *26*, 383–392, doi:10.1016/j.bbi.2011.08.007.

84.  Garcia-Gutierrez, E.; Narbad, A.; Rodríguez, J.M. Autism Spectrum Disorder Associated With Gut Microbiota at Immune, Metabolomic, and Neuroactive Level. *Front. Neurosci.* **2020**, *14*, 1–14, doi:10.3389/fnins.2020.578666.

85.  Moradi, K.; Ashraf-Ganjouei, A.; Tavolinejad, H.; Bagheri, S.; Akhondzadeh, S. The interplay between gut microbiota and autism spectrum disorders: A focus on immunological pathways. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2021**, *106*, 110091, doi:10.1016/j.pnpbp.2020.110091.

86.  Cryan, J.F.; O'riordan, K.J.; Cowan, C.S.M.; Sandhu, K. V.; Bastiaanssen, T.F.S.; Boehme, M.; Codagnone, M.G.; Cussotto, S.; Fulling, C.; Golubeva, A. V.; et al. The microbiota-gut-brain axis. *Physiol. Rev.* **2019**, *99*, 1877–2013, doi:10.1152/physrev.00018.2018.

87.  Marotta, R.; Risoleo, M.C.; Messina, G.; Parisi, L.; Carotenuto, M.; Vetri, L.; Roccella, M. The neurochemistry of autism. *Brain Sci.* **2020**, *10*, 1–18, doi:10.3390/brainsci10030163.

88.  Gabriele, S.; Sacco, R.; Persico, A.M. Blood serotonin levels in autism spectrum disorder: A systematic review and meta-analysis. *Eur. Neuropsychopharmacol.* **2014**, *24*, 919–929, doi:10.1016/j.euroneuro.2014.02.004.

89.  Oblak, A.; Gibbs, T.T.; Blatt, G.J. Reduced serotonin receptor subtypes in a limbic and a neocortical region in autism. *Autism Res.* **2013**, *6*, 571–583, doi:10.1002/aur.1317.

90.  Beversdorf, D.Q.; Nordgren, R.E.; Bonab, A.A.; Fischman, A.J.; Weise, S.B.; Dougherty, D.D.; Felopulos, G.J.; Zhou, F.C.; Bauman, M.L. 5-HT2 receptor distribution shown by [18F] setoperone PET in high-functioning autistic adults. *J. Neuropsychiatry Clin. Neurosci.* **2012**, *24*, 191–197, doi:10.1176/appi.neuropsych.11080202.

91.  Makkonen, I.; Riikonen, R.; Kokki, H.; Airaksinen, M.M.; Kuikka, J.T. Serotonin and dopamine transporter binding in children with autism determined by SPECT. *Dev. Med. Child Neurol.* **2008**, *50*, 593–597, doi:10.1111/j.1469-8749.2008.03027.x.

92.  Nakamura, K.; Sekine, Y.; Ouchi, Y.; Tsujii, M.; Yoshikawa, E.; Futatsubashi, M.; Tsuchiya, K.J.; Sugihara, G.; Iwata, Y.; Suzuki, K.; et al. Brain serotonin and dopamine transporter bindings in adults with high-functioning autism. *Arch. Gen. Psychiatry* **2010**, *67*, 59–68, doi:10.1001/archgenpsychiatry.2009.137.

93.  Jaiswal, P.; Guhathakurta, S.; Singh, A.S.; Verma, D.; Pandey, M.; Varghese, M.; Sinha, S.; Ghosh, S.; Mohanakumar, K.P.; Rajamma, U. SLC6A4 markers modulate platelet 5-HT level and specific behaviors of autism: A study from an Indian population. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2015**, *56*, 196–206, doi:10.1016/j.pnpbp.2014.09.004.

94.  Wongpaiboonwattana, W.; Plong-on, O.; Hnoonual, A.; Limprasert, P. Significant associations between 5-hydroxytryptaminetransporter-linked promoter region polymorphisms of the serotonin transporter (solute carrier family 6 member 4) gene and Thai patients with autism spectrum disorder. *Med.* **2020**, *36*, 1–5,

doi:http://dx.doi.org/10.1097/MD.0000000000021946 1.

95. Hollander, E.; Soorya, L.; Chaplin, W.; Anagnostou, E.; Taylor, B.P.; Ferretti, C.J.; Wasserman, S.; Swanson, E.; Settipani, C. A double-blind placebo-controlled trial of fluoxetine for repetitive behaviors and global severity in adult autism spectrum disorders. *Am. J. Psychiatry* **2012**, *169*, 292–299, doi:10.1176/appi.ajp.2011.10050764.

96. King, B.H.; Hollander, E.; Sikich, L.; McCracken, J.T.; Scahill, L.; Bregman, J.D.; Donnelly, C.L.; Anagnostou, E.; Dukes, K.; Sullivan, L.; et al. Lack of efficacy of citalopram in children with autism spectrum disorders and high levels of repetitive behavior: Citalopram ineffective in children with autism. *Arch. Gen. Psychiatry* **2009**, *66*, 583–590, doi:10.1001/archgenpsychiatry.2009.30.

97. Pourhamzeh, M.; Moravej, F.G.; Arabi, M.; Shahriari, E.; Mehrabi, S.; Ward, R.; Ahadi, R.; Joghataei, M.T. The Roles of Serotonin in Neuropsychiatric Disorders. *Cell. Mol. Neurobiol.* **2021**, doi:10.1007/s10571-021-01064-9.

98. Lake, C.R.; Ziegler, M.G.; Murphy, D.L. Increased Norepinephrine Levels and Decreased Dopamine-β-Hydroxylase Activity in Primary Autism. *Arch. Gen. Psychiatry* **1977**, *34*, 553–556, doi:10.1001/archpsyc.1977.01770170063005.

99. Launay, J.M.; Bursztejn, C.; Ferrari, P.; Dreux, C.; Braconnier, A.; Zarifian, E.; Lancrenon, S.; Fermanian, J. Catecholamines metabolism in infantile autism: A controlled study of 22 autistic children. *J. Autism Dev. Disord.* **1987**, *17*, 333–347, doi:10.1007/BF01487064.

100. Ernst, M.; Zametkin, A.J.; Matochik, J.A.; Pascualvaca, D.; Cohen, R.M. Low medial prefrontal dopaminergic activity in autistic children [4]. *Lancet* **1997**, *350*, 638, doi:10.1016/S0140-6736(05)63326-0.

101. Cai, Y.; Xing, L.; Yang, T.; Chai, R.; Wang, J.; Bao, J.; Shen, W.; Ding, S.; Chen, G. The neurodevelopmental role of dopaminergic signaling in neurological disorders. *Neurosci. Lett.* **2021**, *741*, 135540, doi:10.1016/j.neulet.2020.135540.

102. Pavăl, D. A Dopamine Hypothesis of Autism Spectrum Disorder. *Dev. Neurosci.* **2017**, *39*, 355–360, doi:10.1159/000478725.

103. Di, J.; Li, J.; O'Hara, B.; Alberts, I.; Xiong, L.; Li, J.; Li, X. The role of GABAergic neural circuits in the pathogenesis of autism spectrum disorder. *Int. J. Dev. Neurosci.* **2020**, *80*, 73–85, doi:10.1002/jdn.10005.

104. Tick, B.; Bolton, P.; Happé, F.; Rutter, M.; Rijsdijk, F. Heritability of autism spectrum disorders: A meta-analysis of twin studies. *J. Child Psychol. Psychiatry Allied Discip.* **2016**, *57*, 585–595, doi:10.1111/jcpp.12499.

105. Bailey, A.; Palferman, S.; Heavey, L.; Le Couteur, A. Autism: The phenotype in relatives. *J. Autism Dev. Disord.* **1998**, *28*, 369–392, doi:10.1023/A:1026048320785.

106. Smalley, S.L.; Asarnow, R.F.; Spence, M.A. Autism and Genetics: A Decade of Research. *Arch. Gen. Psychiatry* **1988**, *45*, 953–961, doi:10.1001/archpsyc.1988.01800340081013.

107. Folstein, S.; Rutter, M. Infantile Autism: a Genetic Study of 21 Twin Pairs. *J. Child Psychol. Psychiatry* **1977**, *18*, 297–321, doi:10.1111/j.1469-7610.1977.tb00443.x.

108. Bailey, A.; Le Couteur, A.; Gottesman, I.; Bolton, P.; Simonoff, E.; Yuzda, E.; Rutter, M.; Bailey, A. Autism as a strongly genetic disorder evidence from a british twin Study. *Psychol. Med.* **1995**, *25*, 63–77, doi:10.1017/S0033291700028099.

109. Ritvo, E.R.; Freeman, B.J.; Mason-Brothers, A.; Mo, A. Concordance for the syndrome of autism in 40 pairs of afflicted twins. *Am. J. Psychiatry* **1985**, *142*, 74–77, doi:10.1176/ajp.142.1.74.

110. Steffenburg, S.; Gillberg, C.; Hellgren, L.; Andersson, L.; Gillberg, I.C.; Jakobsson, G.; Bohman, M. A Twin Study of Autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatry* **1989**, *30*, 405–416, doi:10.1111/j.1469-7610.1989.tb00254.x.

111. Gaugler, T.; Klei, L.; Sanders, S.J.; Bodea, C.A.; Goldberg, A.P.; Lee, A.B.; Mahajan, M.; Manaa, D.; Pawitan, Y.; Reichert, J.; et al. Most genetic risk for autism resides with common variation. *Nat. Genet.* **2014**, *46*, 881–885, doi:10.1038/ng.3039.

112. De Rubeis, S.; Buxbaum, J.D. Recent Advances in the Genetics of Autism Spectrum Disorder. *Curr. Neurol. Neurosci. Rep.* **2015**, *15*, doi:10.1007/s11910-015-0553-1.

113. Sandin, S.; Lichtenstein, P.; Kuja-Halkola, R.; Larsson, H.; Hultman, C.M.; Reichenberg, A. The familial risk of autism. *JAMA - J. Am. Med. Assoc.* **2014**, *311*, 1770–1777, doi:10.1001/jama.2014.4144.

114. Klei, L.; Sanders, S.J.; Murtha, M.T.; Hus, V.; Lowe, J.K.; Willsey, A.J.; Moreno-De-Luca, D.; Yu, T.W.; Fombonne, E.; Geschwind, D.; et al. Common genetic variants, acting additively, are a major source of risk for autism. *Mol. Autism* **2012**, *3*, doi:10.1186/2040-2392-3-9.

115. Bourgeron, T. Current knowledge on the genetics of autism and propositions for future research. *Comptes Rendus - Biol.* **2016**, *339*, 300–307, doi:10.1016/j.crvi.2016.05.004.

116. Gillberg, C.; Wahlström, J. Chromosome abnormalities in infantile autism and other childhood psychoses: a population study of 66 cases. *Dev. Med. Child Neurol.* **1985**, *27*, 293–304, doi:10.1111/j.1469-8749.1985.tb04539.x.

117. International molecular genetic study of autism Consortium A full genome screen for autism with evidence for linkage to a region on chromosome 7q. *Hum. Mol. Genet.* **1998**, *7*, 571–578, doi:10.1093/hmg/7.3.571.

118. International molecular genetic study of autism Consortium A genomewide screen for autism: Strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am. J. Hum. Genet.* **2001**, *69*, 570–581, doi:10.1086/323264.

119. S, B.; JC, B.; R, B.; E, B.; TA, B.; TL, C.; D, C.; SE, F.; M, G.; MB, G.; et al. An autosomal genomic screen for autism. *Am. J. Med. Genet.* **1999**, *88*, 609–615, doi:10.1002/(sici)1096-8628(19991215)88:6<609::aid-ajmg7>3.3.co;2-c.

120. Buxbaum, J.D.; Silverman, J.M.; Smith, C.J.; Kilifarski, M.; Reichert, J.; Hollander, E.; Lawlor, B.A.; Fitzgerald, M.; Greenberg, D.A.; Davis, K.L. Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am. J. Hum. Genet.* **2001**, *68*, 1514–1520, doi:10.1086/320588.

121. Liu, J.; Nyholt, D.R.; Magnussen, P.; Parano, E.; Pavone, P.; Geschwind, D.; Lord, C.; Iversen, P.; Hoh, J.; Ott, J.; et al. A genomewide screen for autism susceptibility loci. *Am. J. Hum. Genet.* **2001**, *69*, 327–340, doi:10.1086/321980.

122. Shao, Y.; Cuccaro, M.L.; Hauser, E.R.; Raiford, K.L.; Menold, M.M.; Wolpert, C.M.; Ravan, S.A.; Elston, L.; Decena, K.; Donnelly, S.L.; et al. Fine mapping of autistic disorder to chromosome 15q11-q13 by use of phenotypic subtypes. *Am. J. Hum. Genet.* **2003**, *72*, 539–548, doi:10.1086/367846.

123. Auranen, M.; Vanhala, R.; Varilo, T.; Ayers, K.; Kempas, E.; Ylisaukko-oja, T.; Sinsheimer, J.S.; Peltonen, L.; Järvelä, I. A genomewide screen for autism-spectrum disorders: evidence for a major susceptibility locus on chromosome 3q25-27. *Am. J. Hum. Genet.* **2002**, *71*, 777–790, doi:10.1086/342720.

124. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; Fitzhugh, W.; et al. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921,

doi:10.1038/35057062.

125. Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J.; Sutton, G.G.; Smith, H.O.; Yandell, M.; Evans, C.A.; Holt, R.A.; et al. The sequence of the human genome. *Science (80-. ).* **2001**, *291*, 1304–1351, doi:10.1126/science.1058040.

126. Persico, A.M.; D'agruma, L.; Maiorano, N.; Totaro, A.; Militerni, R.; Bravaccio, C.; Wassink, T.H.; Schneider, C.; Melmed, R.; Trillo, S.; et al. Reelin gene alleles and haplotypes as a factor predisposing to autistic disorder. *Mol. Psychiatry* **2001**, *6*, 150–159, doi:10.1038/sj.mp.4000850.

127. Stromme, P.; Mangelsdorf, M.E.; Scheffer, I.E.; Gécz, J. Infantile spasms, dystonia, and other X-linked phenotypes caused by mutations in Aristaless related homeobox gene, ARX. *Brain Dev.* **2002**, *24*, 266–268, doi:10.1016/S0387-7604(02)00079-7.

128. Jamain, S.; Quach, H.; Betancur, C.; Råstam, M.; Colineaux, C.; Gillberg, C.; Soderstrom, H.; Giros, B.; Leboyer, M.; Gillberg, C.; et al. Mutations of the X-linked genes encoding neuroligins NLGN3 and NLGN4 are associated with autism. *Nat. Genet.* **2003**, *34*, 27–29, doi:10.1038/ng1136.

129. Serajee, F.J.; Nabi, R.; Zhong, H.; Mahbubul Huq, A.H. Association of INPP1, PIK3CG, and TSC2 gene variants with autistic disorder: implications for phosphatidylinositol signalling in autism. *J. Med. Genet.* **2003**, *40*, doi:10.1136/jmg.40.11.e119.

130. Jiang, Y.H.; Sahoo, T.; Michaelis, R.C.; Bercovich, D.; Bressler, J.; Kashork, C.D.; Liu, Q.; Shaffer, L.G.; Schroer, R.J.; Stockton, D.W.; et al. A mixed epigenetic/genetic model for oligogenic inheritance of autism with a limited role for UBE3A. *Am. J. Med. Genet.* **2004**, *131 A*, 1–10, doi:10.1002/ajmg.a.30297.

131. Carney, R.M.; Wolpert, C.M.; Ravan, S.A.; Shahbazian, M.; Ashley-Koch, A.; Cuccaro, M.L.; Vance, J.M.; Pericak-Vance, M.A. Identification of MeCP2 mutations in a series of females with autistic disorder. *Pediatr. Neurol.* **2003**, *28*, 205–211, doi:10.1016/S0887-8994(02)00624-0.

132. Iafrate, A.J.; Feuk, L.; Rivera, M.N.; Listewnik, M.L.; Donahoe, P.K.; Qi, Y.; Scherer, S.W.; Lee, C. Detection of large-scale variation in the human genome. *Nat. Genet.* **2004**, *36*, 949–951, doi:10.1038/ng1416.

133. Sebat, J.; Lakshmi, B.; Troge, J.; Alexander, J.; Young, J.; Lundin, P.; Månér, S.; Massa, H.; Walker, M.; Chi, M.; et al. Large-scale copy number polymorphism in the human genome. *Science (80-. ).* **2004**, *305*, 525–528, doi:10.1126/science.1098918.

134. Sebat, J.; Lakshmi, B.; Malhotra, D.; Troge, J.; Lese-Martin, C.; Walsh, T.; Yamrom, B.; Yoon, S.; Krasnitz, A.; Kendall, J.; et al. Strong association of de novo copy number mutations with autism. *Science (80-. ).* **2007**, *316*, 445–449, doi:10.1126/science.1138659.

135. Bucan, M.; Abrahams, B.S.; Wang, K.; Glessner, J.T.; Herman, E.I.; Sonnenblick, L.I.; Alvarez Retuerto, A.I.; Imielinski, M.; Hadley, D.; Bradfield, J.P.; et al. Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet.* **2009**, *5*, doi:10.1371/journal.pgen.1000536.

136. Levy, D.; Ronemus, M.; Yamrom, B.; Lee, Y. ha; Leotta, A.; Kendall, J.; Marks, S.; Lakshmi, B.; Pai, D.; Ye, K.; et al. Rare De Novo and Transmitted Copy-Number Variation in Autistic Spectrum Disorders. *Neuron* **2011**, *70*, 886–897, doi:10.1016/j.neuron.2011.05.015.

137. Marshall, C.R.; Noor, A.; Vincent, J.B.; Lionel, A.C.; Feuk, L.; Skaug, J.; Shago, M.; Moessner, R.; Pinto, D.; Ren, Y.; et al. Structural Variation of Chromosomes in Autism Spectrum Disorder. *Am. J. Hum. Genet.* **2008**, *82*, 477–

488, doi:10.1016/j.ajhg.2007.12.009.

138. Pinto, D.; Pagnamenta, A.T.; Klei, L.; Anney, R.; Merico, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **2010**, *466*, 368–372, doi:10.1038/nature09146.

139. Sanders, S.J.; Ercan-Sencicek, A.G.; Hus, V.; Luo, R.; Murtha, M.T.; Moreno-De-Luca, D.; Chu, S.H.; Moreau, M.P.; Gupta, A.R.; Thomson, S.A.; et al. Multiple Recurrent De Novo CNVs, Including Duplications of the 7q11.23 Williams Syndrome Region, Are Strongly Associated with Autism. *Neuron* **2011**, *70*, 863–885, doi:10.1016/j.neuron.2011.05.002.

140. Weiss, L.A.; Shen, Y.; Korn, J.M.; Arking, D.E.; Miller, D.T.; Fossdal, R.; Saemundsen, E.; Stefansson, H.; Ferreira, M.A.R.; Green, T.; et al. Association between Microdeletion and Microduplication at 16p11.2 and Autism. *N. Engl. J. Med.* **2008**, *358*, 667–675, doi:10.1056/nejmoa075974.

141. Pizzo, L.; Jensen, M.; Polyak, A.; Rosenfeld, J.A.; Mannik, K.; Krishnan, A.; McCready, E.; Pichon, O.; Le Caignec, C.; Van Dijck, A.; et al. Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants. *Genet. Med.* **2019**, *21*, 816–825, doi:10.1038/s41436-018-0266-3.

142. Sanders, S.J.; He, X.; Willsey, A.J.; Ercan-Sencicek, A.G.; Samocha, K.E.; Cicek, A.E.; Murtha, M.T.; Bal, V.H.; Bishop, S.L.; Dong, S.; et al. Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* **2015**, *87*, 1215–1233, doi:10.1016/j.neuron.2015.09.016.

143. Vicari, S.; Napoli, E.; Cordeddu, V.; Menghini, D.; Alesi, V.; Loddo, S.; Novelli, A.; Tartaglia, M. Copy number variants in autism spectrum disorders. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2019**, *92*, 421–427, doi:10.1016/j.pnpbp.2019.02.012.

144. Brandler, W.M.; Sebat, J. From de novo mutations to personalized therapeutic interventions in autism. *Annu. Rev. Med.* **2015**, *66*, 487–507, doi:10.1146/annurev-med-091113-024550.

145. Pinto, D.; Delaby, E.; Merico, D.; Barbosa, M.; Merikangas, A.; Klei, L.; Thiruvahindrapuram, B.; Xu, X.; Ziman, R.; Wang, Z.; et al. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am. J. Hum. Genet.* **2014**, *94*, 677–694, doi:10.1016/j.ajhg.2014.03.018.

146. Iossifov, I.; Ronemus, M.; Levy, D.; Wang, Z.; Hakker, I.; Rosenbaum, J.; Yamrom, B.; Lee, Y. ha; Narzisi, G.; Leotta, A.; et al. De Novo Gene Disruptions in Children on the Autistic Spectrum. *Neuron* **2012**, *74*, 285–299, doi:10.1016/j.neuron.2012.04.009.

147. Neale, B.M.; Kou, Y.; Liu, L.; Ma'ayan, A.; Samocha, K.E.; Sabo, A.; Lin, C.-F.; Stevens, C.; Wang, L.-S.; Makarov, V.; et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **2012**, *485*, 242–246, doi:10.1038/nature11011.

148. O'Roak, B.J.; Vives, L.; Girirajan, S.; Karakoc, E.; Krumm, N.; Coe, B.P.; Levy, R.; Ko, A.; Lee, C.; Smith, J.D.; et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **2012**, *485*, 246–250, doi:10.1038/nature10989.

149. Sanders, S.J.; Murtha, M.T.; Gupta, A.R.; Murdoch, J.D.; Raubeson, M.J.; Willsey, A.J.; Ercan-Sencicek, A.G.; Di Lullo, N.M.; Parikshak, N.N.; Stein, J.L.; et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **2012**, *485*, 237–241, doi:10.1038/nature10945.

150. Chahrour, M.H.; Yu, T.W.; Lim, E.T.; Ataman, B.; Coulter, M.E.; Hill, R.S.; Stevens, C.R.; Schubert, C.R.; Greenberg,

M.E.; Gabriel, S.B.; et al. Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* **2012**, *8*, doi:10.1371/journal.pgen.1002635.

151. Lim, E.T.; Raychaudhuri, S.; Sanders, S.J.; Stevens, C.; Sabo, A.; MacArthur, D.G.; Neale, B.M.; Kirby, A.; Ruderfer, D.M.; Fromer, M.; et al. Rare Complete Knockouts in Humans: Population Distribution and Significant Role in Autism Spectrum Disorders. *Neuron* **2013**, *77*, 235–242, doi:10.1016/j.neuron.2012.12.029.

152. Iossifov, I.; Levy, D.; Allen, J.; Ye, K.; Ronemus, M.; Lee, Y.H.; Yamrom, B.; Wigler, M. Low load for disruptive mutations in autism genes and their biased transmission. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E5600–E5607, doi:10.1073/pnas.1516376112.

153. Krumm, N.; Turner, T.N.; Baker, C.; Vives, L.; Mohajeri, K.; Witherspoon, K.; Raja, A.; Coe, B.P.; Stessman, H.A.; He, Z.X.; et al. Excess of rare, inherited truncating mutations in autism. *Nat. Genet.* **2015**, *47*, 582–588, doi:10.1038/ng.3303.

154. Morrow, E.M.; Yoo, S.Y.; Flavell, S.W.; Kim, T.K.; Lin, Y.; Hill, R.S.; Mukaddes, N.M.; Balkhy, S.; Gascon, G.; Hashmi, A.; et al. Identifying autism loci and genes by tracing recent shared ancestry. *Science (80-. ).* **2008**, *321*, 218–223, doi:10.1126/science.1157657.

155. Yu, T.W.; Chahrour, M.H.; Coulter, M.E.; Jiralerspong, S.; Okamura-Ikeda, K.; Ataman, B.; Schmitz-Abe, K.; Harmin, D.A.; Adli, M.; Malik, A.N.; et al. Using Whole-Exome Sequencing to Identify Inherited Causes of Autism. *Neuron* **2013**, *77*, 259–273, doi:10.1016/j.neuron.2012.11.002.

156. Iossifov, I.; O'Roak, B.J.; Sanders, S.J.; Ronemus, M.; Krumm, N.; Levy, D.; Stessman, H.A.; Witherspoon, K.T.; Vives, L.; Patterson, K.E.; et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **2014**, *515*, 216–221, doi:10.1038/nature13908.

157. Robinson, E.B.; Samocha, K.E.; Kosmicki, J.A.; McGrath, L.; Neale, B.M.; Perlis, R.H.; Daly, M.J. Autism spectrum disorder severity reflects the average contribution of de novo and familial influences. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 15161–15165, doi:10.1073/pnas.1409204111.

158. Girirajan, S.; Brkanac, Z.; Coe, B.P.; Baker, C.; Vives, L.; Vu, T.H.; Shafer, N.; Bernier, R.; Ferrero, G.B.; Silengo, M.; et al. Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet.* **2011**, *7*, doi:10.1371/journal.pgen.1002334.

159. Sanders, S.J. Next-generation sequencing in autism spectrum disorder. *Cold Spring Harb. Perspect. Med.* **2019**, *9*, doi:10.1101/cshperspect.a026872.

160. Brandler, W.M.; Antaki, D.; Gujral, M.; Noor, A.; Rosanio, G.; Chapman, T.R.; Barrera, D.J.; Lin, G.N.; Malhotra, D.; Watts, A.C.; et al. Frequency and Complexity of de Novo Structural Mutation in Autism. *Am. J. Hum. Genet.* **2016**, *98*, 667–679, doi:10.1016/j.ajhg.2016.02.018.

161. Collins, R.L.; Brand, H.; Redin, C.E.; Hanscom, C.; Antolik, C.; Stone, M.R.; Glessner, J.T.; Mason, T.; Pregno, G.; Dorrani, N.; et al. Defining the diverse spectrum of inversions, complex structural variation, and chromothripsis in the morbid human genome. *Genome Biol.* **2017**, *18*, doi:10.1186/s13059-017-1158-6.

162. Werling, D.M.; Brand, H.; An, J.Y.; Stone, M.R.; Zhu, L.; Glessner, J.T.; Collins, R.L.; Dong, S.; Layer, R.M.; Markenscoff-Papadimitriou, E.; et al. An analytical framework for whole-genome sequence association studies and its implications for autism spectrum disorder. *Nat. Genet.* **2018**, *50*, 727–736, doi:10.1038/s41588-018-0107-y.

163. Turner, T.N.; Hormozdiari, F.; Duyzend, M.H.; McClymont, S.A.; Hook, P.W.; Iossifov, I.; Raja, A.; Baker, C.;

Hoekzema, K.; Stessman, H.A.; et al. Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. *Am. J. Hum. Genet.* **2016**, *98*, 58–74, doi:10.1016/j.ajhg.2015.11.023.

164. Yuen, R.K.C.; Merico, D.; Bookman, M.; Howe, J.L.; Thiruvahindrapuram, B.; Patel, R. V.; Whitney, J.; Deflaux, N.; Bingham, J.; Wang, Z.; et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat. Neurosci.* **2017**, *20*, 602–611, doi:10.1038/nn.4524.

165. De Rubeis, S.; He, X.; Goldberg, A.P.; Poultney, C.S.; Samocha, K.; Cicek, A.E.; Kou, Y.; Liu, L.; Fromer, M.; Walker, S.; et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **2014**, *515*, 209–215, doi:10.1038/nature13772.

166. Lauritsen, M.B.; Als, T.D.; Dahl, H.A.; Flint, T.J.; Wang, A.G.; Vang, M.; Kruse, T.A.; Ewald, H.; Mors, O. A genome-wide search for alleles and haplotypes associated with autism and related pervasive developmental disorders on the Faroe Islands. *Mol. Psychiatry* **2006**, *11*, 37–46, doi:10.1038/sj.mp.4001754.

167. Arking, D.E.; Cutler, D.J.; Brune, C.W.; Teslovich, T.M.; West, K.; Ikeda, M.; Rea, A.; Guy, M.; Lin, S.; Cook, E.H.; et al. A Common Genetic Variant in the Neurexin Superfamily Member CNTNAP2 Increases Familial Risk of Autism. *Am. J. Hum. Genet.* **2008**, *82*, 160–164, doi:10.1016/j.ajhg.2007.09.015.

168. Anney, R.; Klei, L.; Pinto, D.; Almeida, J.; Bacchelli, E.; Baird, G.; Bolshakova, N.; Bölte, S.; Bolton, P.F.; Bourgeron, T.; et al. Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum. Mol. Genet.* **2012**, *21*, 4781–4792, doi:10.1093/hmg/dds301.

169. Wang, K.; Zhang, H.; Ma, D.; Bucan, M.; Glessner, J.T.; Abrahams, B.S.; Salyakina, D.; Imielinski, M.; Bradfield, J.P.; Sleiman, P.M.A.; et al. Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature* **2009**, *459*, 528–533, doi:10.1038/nature07999.

170. Ma, D.; Salyakina, D.; Jaworski, J.M.; Konidari, I.; Whitehead, P.L.; Andersen, A.N.; Hoffman, J.D.; Slifer, S.H.; Hedges, D.J.; Cukier, H.N.; et al. A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *Ann. Hum. Genet.* **2009**, *73*, 263–273, doi:10.1111/j.1469-1809.2009.00523.x.

171. Weiss, L.A.; Arking, D.E.; Daly, M.J.; Chakravarti, A.; Brune, C.W.; West, K.; O'Connor, A.; Hilton, G.; Tomlinson, R.L.; West, A.B.; et al. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **2009**, *461*, 802–808, doi:10.1038/nature08490.

172. Anney, R.; Klei, L.; Pinto, D.; Regan, R.; Conroy, J.; Magalhaes, T.R.; Correia, C.; Abrahams, B.S.; Sykes, N.; Pagnamenta, A.T.; et al. A genome-wide scan for common alleles affecting risk for autism. *Hum. Mol. Genet.* **2010**, *19*, 4072–4082, doi:10.1093/hmg/ddq307.

173. Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **2017**, *8*, 21, doi:10.1186/s13229-017-0137-9.

174. Grove, J.; Ripke, S.; Als, T.D.; Mattheisen, M.; Walters, R.K.; Won, H.; Pallesen, J.; Agerbo, E.; Andreassen, O.A.; Anney, R.; et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **2019**, *51*, 431–444, doi:10.1038/s41588-019-0344-8.

175. Rylaarsdam, L.; Guemez-Gamboa, A. Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front. Cell. Neurosci.* **2019**, *13*, 1–15, doi:10.3389/fncel.2019.00385.

176. Wiśniowiecka-Kowalnik, B.; Nowakowska, B.A. Genetics and epigenetics of autism spectrum disorder—current evidence in the field. *J. Appl. Genet.* **2019**, *60*, 37–47, doi:10.1007/s13353-018-00480-w.

177. Modabbernia, A.; Velthorst, E.; Reichenberg, A. Environmental risk factors for autism: an evidence-based review of systematic reviews and meta-analyses. *Mol. Autism* 2017, *8*.

178. Wu, S.; Wu, F.; Ding, Y.; Hou, J.; Bi, J.; Zhang, Z. Advanced parental age and autism risk in children: a systematic review and meta-analysis. *Acta Psychiatr. Scand.* **2017**, *135*, 29–41, doi:10.1111/acps.12666.

179. Taylor, L.E.; Swerdfeger, A.L.; Eslick, G.D. Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine* **2014**, *32*, 3623–3629, doi:10.1016/j.vaccine.2014.04.085.

180. Iakoucheva, L.M.; Muotri, A.R.; Sebat, J. Getting to the Cores of Autism. *Cell* **2019**, *178*, 1287–1298, doi:10.1016/j.cell.2019.07.037.

181. Toma, C. Genetic Variation across Phenotypic Severity of Autism. *Trends Genet.* **2020**, 1639, doi:10.1016/j.tig.2020.01.005.

182. Bourgeron, T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. *Nat. Rev. Neurosci.* **2015**, *16*, 551–563, doi:10.1038/nrn3992.

183. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A.; et al. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753, doi:10.1038/nature08494.

184. Weiner, D.J.; Wigdor, E.M.; Ripke, S.; Walters, R.K.; Kosmicki, J.A.; Grove, J.; Samocha, K.E.; Goldstein, J.I.; Okbay, A.; Bybjerg-Grauholm, J.; et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat. Genet.* **2017**, *49*, 978–985, doi:10.1038/ng.3863.

185. Niemi, M.E.K.; Martin, H.C.; Rice, D.L.; Gallone, G.; Gordon, S.; Kelemen, M.; McAloney, K.; McRae, J.; Radford, E.J.; Yu, S.; et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* **2018**, *562*, 268–271, doi:10.1038/s41586-018-0566-4.

186. Middeldorp, C.M.; Wray, N.R. The value of polygenic analyses in psychiatry. *World Psychiatry* **2018**, *17*, 26–28, doi:DOI:10.1002/wps.20480.

187. Boyle, E.A.; Li, Y.I.; Pritchard, J.K. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* **2017**, *169*, 1177–1186, doi:10.1016/j.cell.2017.05.038.

188. Liu, X.; Li, Y.I.; Pritchard, J.K. Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell* **2019**, *177*, 1022-1034.e6, doi:10.1016/j.cell.2019.04.014.

189. Wray, N.R.; Wijmenga, C.; Sullivan, P.F.; Yang, J.; Visscher, P.M. Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. *Cell* **2018**, *173*, 1573–1580, doi:10.1016/j.cell.2018.05.051.

190. Solovieff, N.; Cotsapas, C.; Lee, P.H.; Purcell, S.M.; Smoller, J.W. Pleiotropy in complex traits: Challenges and strategies. *Nat. Rev. Genet.* **2013**, *14*, 483–495, doi:10.1038/nrg3461.

191. Lee, P.H.; Feng, Y.C.A.; Smoller, J.W. Pleiotropy and Cross-Disorder Genetics Among Psychiatric Disorders. *Biol. Psychiatry* **2021**, *89*, 20–31, doi:10.1016/j.biopsych.2020.09.026.

192. Lee, P.H.; Anttila, V.; Erneri; Won, H.; Feng, Y.-C.A.; Rosenthal, J.; Zhu, Z.; Tucker-Drob, E.M.; Nivard, M.G.; Grotzinger, A.D.; et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **2019**, *179*, 1469-1482.e11, doi:10.1016/j.cell.2019.11.020.

193. Grotzinger, A.D.; Mallard, T.T.; Akingbuwa, W.A.; Ip, H.F.; Adams, M.J.; Lewis, C.M.; McIntosh, A.M.; Grove, J.; Dalsgaard, S.; Lesch, K.-P.; et al. Genetic Architecture of 11 Major Psychiatric Disorders at Biobehavioral,

Functional 2 Genomic, and Molecular Genetic Levels of Analysis. *medRxiv* **2020**, doi:https://doi.org/10.1101/2020.09.22.20196089.

194. Ma, S.L.; Chen, L.H.; Lee, C.C.; Lai, K.Y.C.; Hung, S.F.; Tang, C.P.; Ho, T.P.; Shea, C.; Mo, F.; Mak, T.S.H.; et al. Genetic Overlap Between Attention Deficit/Hyperactivity Disorder and Autism Spectrum Disorder in SHANK2 Gene. *Front. Neurosci.* **2021**, *15*, doi:10.3389/fnins.2021.649588.

195. Carroll, L.S.; Owen, M.J. Genetic overlap between autism, schizophrenia and bipolar disorder. *Genome Med.* **2009**, *1*, 1–7.

196. Constantino, J.N.; Todd, R.D. Autistic traits in the general population: A twin study. *Arch. Gen. Psychiatry* **2003**, *60*, 524–530, doi:10.1001/archpsyc.60.5.524.

197. Robinson, E.B.; Munir, K.; Munaf, M.R.; Hughes, M.; McCormick, M.C.; Koenen, K.C. Stability of autistic traits in the general population: Further evidence for a continuum of impairment. *J. Am. Acad. Child Adolesc. Psychiatry* **2011**, *50*, 376–384, doi:10.1016/j.jaac.2011.01.005.

198. Robinson, E.B.; St Pourcain, B.; Anttila, V.; Kosmicki, J.A.; Bulik-Sullivan, B.; Grove, J.; Maller, J.; Samocha, K.E.; Sanders, S.J.; Ripke, S.; et al. Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nat. Genet.* **2016**, *48*, 552–555, doi:10.1038/ng.3529.

199. Wakabayashi, A.; Baron-Cohen, S.; Wheelwright, S.; Tojo, Y. The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. *J. Autism Dev. Disord.* **2006**, *36*, 263–270, doi:10.1007/s10803-005-0061-2.

200. Constantino, J.N.; Zhang, Y.; Frazier, T.; Abbacchi, A.M.; Law, P. Sibling recurrence and the genetic epidemiology of autism. *Am. J. Psychiatry* **2010**, *167*, 1349–1356, doi:10.1176/appi.ajp.2010.09101470.

201. McIntosh, A.M.; Sullivan, P.F.; Lewis, C.M. Uncovering the Genetic Architecture of Major Depression. *Neuron* 2019, *102*, 91–103.

202. Clarke, T.K.; Lupton, M.K.; Fernandez-Pujals, A.M.; Starr, J.; Davies, G.; Cox, S.; Pattie, A.; Liewald, D.C.; Hall, L.S.; Macintyre, D.J.; et al. Common polygenic risk for autism spectrum disorder (ASD) is associated with cognitive ability in the general population. *Mol. Psychiatry* **2016**, *21*, 419–425, doi:10.1038/mp.2015.12.

203. Schork, A.J.; Brown, T.T.; Hagler, D.J.; Thompson, W.K.; Chen, C.H.; Dale, A.M.; Jernigan, T.L.; Akshoomoff, N. Polygenic risk for psychiatric disorders correlates with executive function in typical development. *Genes, Brain Behav.* **2019**, *18*, 1–9, doi:10.1111/gbb.12480.

204. Khundrakpam, B.; Vainik, U.; Gong, J.; Al-Sharif, N.; Bhutani, N.; Kiar, G.; Zeighami, Y.; Kirschner, M.; Luo, C.; Dagher, A.; et al. Neural correlates of polygenic risk score for autism spectrum disorders in general population. *Brain Commun.* **2020**, *2*, 1–13, doi:10.1093/braincomms/fcaa092.

205. Satterstrom, F.K.; Kosmicki, J.A.; Wang, J.; Breen, M.S.; De Rubeis, S.; An, J.-Y.; Peng, M.; Collins, R.; Grove, J.; Klei, L.; et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* **2020**, *180*, 568–584, doi:10.1101/484113.

206. Ayhan, F.; Konopka, G. Regulatory genes and pathways disrupted in autism spectrum disorders. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2019**, *89*, 57–64, doi:10.1016/j.pnpbp.2018.08.017.

207. Aghazadeh, Y.; Papadopoulos, V. The role of the 14-3-3 protein family in health, disease, and drug development. *Drug Discov. Today* **2016**, *21*, 278–287, doi:10.1016/j.drudis.2015.09.012.

208. Cornell, B.; Toyo-oka, K. 14-3-3 proteins in brain development: Neurogenesis, neuronal migration and neuromorphogenesis. *Front. Mol. Neurosci.* **2017**, *10*, 318, doi:10.3389/fnmol.2017.00318.

209. Hermeking, H.; Benzinger, A. 14-3-3 Proteins in Cell Cycle Regulation. *Semin. Cancer Biol.* **2006**, *16*, 183–192, doi:10.1016/j.semcancer.2006.03.002.

210. Kaplan, A.; Kent, C.B.; Charron, F.; Fournier, A.E. Switching responses: Spatial and temporal regulators of axon guidance. *Mol. Neurobiol.* **2014**, *49*, 1077–1086, doi:10.1007/s12035-013-8582-8.

211. Ikeda, M.; Hikita, T.; Taya, S.; Uraguchi-asaki, J.; Toyo-Oka, K.; Wynshaw-boris, A.; Ujike, H.; Inada, T.; Takao, K.; Miyakawa, T.T.; et al. Identification of YWHAE, a gene encoding 14-3-3epsilon, as a possible susceptibility gene for schizophrenia. *Hum. Mol. Genet.* **2008**, *17*, 3212–3222, doi:10.1093/hmg/ddn217.

212. Jia, Y.; Yu, X.; Zhang, B.; Yuan, Y.; Xu, Q.; Shen, Y.; Shen, Y. An association study between polymorphisms in three genes of 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein) family and paranoid schizophrenia in northern Chinese population. *Eur. Psychiatry* **2004**, *19*, 377–379, doi:10.1016/j.eurpsy.2004.07.006.

213. Li, Z.; Chen, J.; Yu, H.; He, L.; Xu, Y.; Zhang, D.; Yi, Q.; Li, C.; Li, X.; Shen, J.; et al. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat. Genet.* **2017**, *49*, 1576–1583, doi:10.1038/ng.3973.

214. Oldmeadow, C.; Mossman, D.; Evans, T.J.; Holliday, E.G.; Tooney, P.A.; Cairns, M.J.; Wu, J.; Carr, V.; Attia, J.R.; Scott, R.J. Combined analysis of exon splicing and genome wide polymorphism data predict schizophrenia risk loci. *J. Psychiatr. Res.* **2014**, *52*, 44–49, doi:10.1016/j.jpsychires.2014.01.011.

215. Wong, A.H.C.; Macciardi, F.; Klempan, T.; Kawczynski, W.; Barr, C.L.; Lakatoo, S.; Wong, M.; Buckle, C.; Trakalo, J.; Boffa, E.; et al. Identification of candidate genes for psychosis in rat models, and possible association between schizophrenia and the 14-3-3η gene. *Mol. Psychiatry* **2003**, *8*, 156–166, doi:10.1038/sj.mp.4001237.

216. Jacobsen, K.K.; Kleppe, R.; Johansson, S.; Zayats, T.; Haavik, J. Epistatic and gene wide effects in YWHA and aromatic amino hydroxylase genes across ADHD and other common neuropsychiatric disorders: Association with YWHAE. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2015**, *168*, 423–432, doi:10.1002/ajmg.b.32339.

217. Grover, D.; Verma, R.; Goes, F.S.; Belmonte Mahon, P.L.; Gershon, E.S.; McMahon, F.J.; Potash, J.B.; Steele, J.; Kassem, L.; Lopez, V.; et al. Family-based association of YWHAH in psychotic bipolar disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2009**, *150*, 977–983, doi:10.1002/ajmg.b.30927.

218. Liu, J.; Zhang, H.X.; Li, Z.Q.; Li, T.; Li, J.Y.; Wang, T.; Li, Y.; Feng, G.Y.; Shi, Y.Y.; He, L. The YWHAE gene confers risk to major depressive disorder in the male group of Chinese Han population. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2017**, *77*, 172–177, doi:10.1016/j.pnpbp.2017.04.013.

219. Yanagi, M.; Shirakawa, O.; Kitamura, N.; Okamura, K.; Sakurai, K.; Nishiguchi, N.; Hashimoto, T.; Nushida, H.; Ueno, Y.; Kanbe, D.; et al. Association of 14-3-3 ε gene haplotype with completed suicide in Japanese. *J. Hum. Genet.* **2005**, *50*, 210–216, doi:10.1007/s10038-005-0241-0.

220. Cheah, P.S.; Ramshaw, H.S.; Thomas, P.Q.; Toyo-Oka, K.; Xu, X.; Martin, S.; Coyle, P.; Guthridge, M.A.; Stomski, F.; Van Den Buuse, M.; et al. Neurodevelopmental and neuropsychiatric behaviour defects arise from 14-3-3ζ deficiency. *Mol. Psychiatry* **2012**, *17*, 451–466, doi:10.1038/mp.2011.158.

221. Toyo-Oka, K.; Shionoya, A.; Gambello, M.J.; Cardoso, C.; Leventer, R.; Ward, H.L.; Ayala, R.; Tsai, L.H.; Dobyns, W.; Ledbetter, D.; et al. 14-3-3ε is important for neuronal migration by binding to NUDEL: A molecular explanation for Miller-Dieker syndrome. *Nat. Genet.* **2003**, *34*, 274–285, doi:10.1038/ng1169.

222. Xu, X.; Jaehne, E.J.; Greenberg, Z.; McCarthy, P.; Saleh, E.; Parish, C.L.; Camera, D.; Heng, J.; Haas, M.; Baune, B.T.; et al. 14-3-3ζ deficient mice in the BALB/c background display behavioural and anatomical defects

associated with neurodevelopmental disorders. *Sci. Rep.* **2015**, *5*, doi:10.1038/srep12434.

223. Wachi, T.; Cornell, B.; Toyo-oka, K. Complete ablation of the 14-3-3epsilon protein results in multiple defects in neuropsychiatric behaviors. *Behav. Brain Res.* **2017**, *319*, 31–36, doi:10.1016/j.bbr.2016.11.016.

224. Kim, D.E.; Cho, C.H.; Sim, K.M.; Kwon, O.; Hwang, E.M.; Kim, H.W.; Park, J.Y. 14-3-3Γ Haploinsufficient Mice Display Hyperactive and Stress-Sensitive Behaviors. *Exp. Neurobiol.* **2019**, *28*, 43–53, doi:10.5607/en.2019.28.1.43.

225. Graham, K.; Zhang, J.; Qiao, H.; Wu, Y.; Zhou, Y. Region-specific inhibition of 14-3-3 proteins induces psychomotor behaviors in mice. *npj Schizophr.* **2019**, *5*, 1, doi:10.1038/s41537-018-0069-1.

226. Toyo-Oka, K.; Wachi, T.; Hunt, R.F.; Baraban, S.C.; Taya, S.; Ramshaw, H.; Kaibuchi, K.; Schwarz, Q.P.; Lopez, A.F.; Wynshaw-Boris, A. 14-3-3E and Z Regulate Neurogenesis and Differentiation of Neuronal Progenitor Cells in the Developing Brain. *J. Neurosci.* **2014**, *34*, 12168–12181, doi:10.1523/JNEUROSCI.2513-13.2014.

227. Toma, C.; Torrico, B.; Hervás, A.; Valdés-Mas, R.; Tristán-Noguero, A.; Padillo, V.; Maristany, M.; Salgado, M.; Arenas, C.; Puente, X.S.; et al. Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. *Mol. Psychiatry* **2014**, *19*, 784–790, doi:10.1038/mp.2013.106.

228. Bill, B.R.; Lowe, J.K.; DyBuncio, C.T.; Fogel, B.L. Orchestration of neurodevelopmental programs by RBFOX1: Implications for autism spectrum disorder. In *International Review of Neurobiology*; Academic Press Inc., 2013; Vol. 113, pp. 251–267.

229. Conboy, J.G. Developmental regulation of RNA processing by Rbfox proteins. *Wiley Interdiscip. Rev. RNA* 2017, *8*.

230. Li, Y.I.; Sanchez-Pulido, L.; Haerty, W.; Ponting, C.P. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res.* 2015, *25*, 1–13.

231. Hamada, N.; Ito, H.; Nishijo, T.; Iwamoto, I.; Morishita, R.; Tabata, H.; Momiyama, T.; Nagata, K.I. Essential role of the nuclear isoform of RBFOX1, a candidate gene for autism spectrum disorders, in the brain development. *Sci. Rep.* **2016**, *6*, 1–19, doi:10.1038/srep30805.

232. Gehman, L.T.; Stoilov, P.; Maguire, J.; Damianov, A.; Lin, C.H.; Shiue, L.; Ares, M.; Mody, I.; Black, D.L. The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat. Genet.* **2011**, *43*, 706–711, doi:10.1038/ng.841.

233. Lee, J.A.; Damianov, A.; Lin, C.H.; Fontes, M.; Parikshak, N.N.; Anderson, E.S.; Geschwind, D.H.; Black, D.L.; Martin, K.C. Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. *Neuron* **2016**, *89*, 113–128, doi:10.1016/j.neuron.2015.11.025.

234. Prashad, S.; Gopal, P.P. RNA-binding proteins in neurological development and disease. *RNA Biol.* **2020**, *00*, 1–16, doi:10.1080/15476286.2020.1809186.

235. Bacchelli, E.; Cameli, C.; Viggiano, M.; Igliozzi, R.; Mancini, A.; Tancredi, R.; Battaglia, A.; Maestrini, E. An integrated analysis of rare CNV and exome variation in Autism Spectrum Disorder using the Infinium PsychArray. *Sci. Rep.* **2020**, *10*, 3198, doi:10.1038/s41598-020-59922-3.

236. Griswold, A.J.; Dueker, N.D.; Van Booven, D.; Rantus, J.A.; Jaworski, J.M.; Slifer, S.H.; Schmidt, M.A.; Hulme, W.; Konidari, I.; Whitehead, P.L.; et al. Targeted massively parallel sequencing of autism spectrum disorder-associated genes in a case control cohort reveals rare loss-of-function risk variants. *Mol. Autism* **2015**, *6*, 1–11, doi:10.1186/s13229-015-0034-z.

237. Zhao, W.W. Intragenic deletion of RBFOX1 associated with neurodevelopmental/ neuropsychiatric disorders and possibly other clinical presentations. *Mol. Cytogenet.* **2013**, *6*, 1–5, doi:10.1186/1755-8166-6-26.

238. Voineagu, I.; Wang, X.; Johnston, P.; Lowe, J.K.; Tian, Y.; Horvath, S.; Mill, J.; Cantor, R.M.; Blencowe, B.J.; Geschwind, D.H. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **2011**, *474*, 380–386, doi:10.1038/nature10110.

239. Bhalla, K.; Phillips, H.A.; Crawford, J.; McKenzie, O.L.D.; Mulley, J.C.; Eyre, H.; Gardner, A.E.; Kremmidiotis, G.; Callen, D.F. The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J. Hum. Genet.* **2004**, *49*, 308–311, doi:10.1007/s10038-004-0145-4.

240. Martin, C.L.; Duvall, J.A.; Ilkin, Y.; Simon, J.S.; Arreaza, M.G.; Wilkes, K.; Alvarez-Retuerto, A.; Whichello, A.; Powell, C.M.; Rao, K.; et al. Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2007**, *144*, 869–876, doi:10.1002/ajmg.b.30530.

241. Fernàndez-Castillo, N.; Gan, G.; van Donkelaar, M.M.J.; Vaht, M.; Weber, H.; Retz, W.; Meyer-Lindenberg, A.; Franke, B.; Harro, J.; Reif, A.; et al. RBFOX1, encoding a splicing regulator, is a candidate gene for aggressive behavior. *Eur. Neuropsychopharmacol.* **2020**, *30*, 44–55, doi:10.1016/j.euroneuro.2017.11.012.

242. Kushima, I.; Aleksic, B.; Nakatochi, M.; Shimamura, T.; Okada, T.; Uno, Y.; Morikawa, M.; Ishizuka, K.; Shiino, T.; Kimura, H.; et al. Comparative Analyses of Copy-Number Variation in Autism Spectrum Disorder and Schizophrenia Reveal Etiological Overlap and Biological Insights. *Cell Rep.* **2018**, *24*, 2838–2856, doi:10.1016/j.celrep.2018.08.022.

243. Elia, J.; Gai, X.; Xie, H.M.; Perin, J.C.; Geiger, E.; Glessner, J.T.; D'Arcy, M.; Deberardinis, R.; Frackelton, E.; Kim, C.; et al. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry* **2010**, *15*, 637–646, doi:10.1038/mp.2009.57.

244. Wray, N.R.; Ripke, S.; Mattheisen, M.; Trzaskowski, M.; Byrne, E.M.; Abdellaoui, A.; Adams, M.J.; Agerbo, E.; Air, T.M.; Andlauer, T.M.F.; et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **2018**, *50*, 668–681, doi:10.1038/s41588-018-0090-3.

245. Fogel, B.L.; Wexler, E.; Wahnich, A.; Friedrich, T.; Vijayendran, C.; Gao, F.; Parikshak, N.; Konopka, G.; Geschwind, D.H. RBFOX1 regulates both splicing and transcriptional networks in human neuronal development. *Hum. Mol. Genet.* **2012**, *21*, 4171–4186, doi:10.1093/hmg/dds240.

246. Kazi, J.U.; Kabir, N.N.; Rönnstrand, L. Brain-Expressed X-linked (BEX) proteins in human cancers. *Biochim. Biophys. Acta - Rev. Cancer* **2015**, *1856*, 226–233, doi:10.1016/j.bbcan.2015.09.001.

247. Perrin, F.E.; Boisset, G.; Lathuilière, A.; Kato, A.C. Cell death pathways differ in several mouse models with motoneurone disease: Analysis of pure motoneurone populations at a presymptomatic age. *J. Neurochem.* **2006**, *98*, 1959–1972, doi:10.1111/j.1471-4159.2006.04024.x.

248. Khazaei, M.R.; Halfter, H.; Karimzadeh, F.; Koo, J.H.; Margolis, F.L.; Young, P. Bex1 is involved in the regeneration of axons after injury. *J. Neurochem.* **2010**, *115*, 910–920, doi:10.1111/j.1471-4159.2010.06960.x.

249. Mathys, H.; Davila-Velderrain, J.; Peng, Z.; Gao, F.; Mohammadi, S.; Young, J.Z.; Menon, M.; He, L.; Abdurrob, F.; Jiang, X.; et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **2019**, *570*, 332–337, doi:10.1038/s41586-019-1195-2.

250. Yamamoto, T.; Wilsdon, A.; Joss, S.; Isidor, B.; Erlandsson, A.; Suri, M.; Sangu, N.; Shimada, S.; Shimojima, K.; Le

# REFERENCES

Caignec, C.; et al. An emerging phenotype of Xq22 microdeletions in females with severe intellectual disability, hypotonia and behavioral abnormalities. *J. Hum. Genet.* **2014**, *59*, 300–306, doi:10.1038/jhg.2014.21.

251. Shirai, K.; Higashi, Y.; Shimojima, K.; Yamamoto, T. An Xq22.1q22.2 nullisomy in a male patient with severe neurological impairment. *Am. J. Med. Genet. Part A* **2017**, *173*, 1124–1127, doi:10.1002/ajmg.a.38134.

252. Hijazi, H.; Coelho, F.S.; Gonzaga-Jauregui, C.; Bernardini, L.; Mar, S.S.; Manning, M.A.; Hanson-Kahn, A.; Naidu, S.B.; Srivastava, S.; Lee, J.A.; et al. Xq22 deletions and correlation with distinct neurological disease traits in females: Further evidence for a contiguous gene syndrome. *Hum. Mutat.* **2020**, *41*, 150–168, doi:10.1002/humu.23902.

253. Pensado-López, A.; Veiga-Rúa, S.; Carracedo, Á.; Allegue, C.; Sánchez, L. Experimental models to study autism spectrum disorders: Hipscs, rodents and zebrafish. *Genes (Basel).* **2020**, *11*, 1–45, doi:10.3390/genes11111376.

254. Brust, V.; Schindler, P.M.; Lewejohann, L. Lifetime development of behavioural phenotype in the house mouse (Mus musculus). *Front. Zool.* **2015**, *12 (Suppl1*, S17, doi:10.1186/1742-9994-12-S1-S17.

255. Silverman, J.L.; Yang, M.; Lord, C.; Crawley, J.N. Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* **2010**, *11*, 490–502, doi:10.1038/nrn2851.

256. Kazdoba, T.M.; Prescott, T.L.; Yang, M.; Silverman, J.L.; Solomon, M.; Crawley, J.N. Translational Mouse Models of Autism: Advancing Toward Pharmacological Therapeutics. *Transl. Neuropsychopharmacol. Curr. Top. Behav. Neurosci.* **2016**, *28*, 1–52, doi:10.1007/7854_2015_5003.

257. Chang, Y.C.; Cole, T.B.; Costa, L.G. Behavioral phenotyping for autism spectrum disorders in mice. *Curr. Protoc. Toxicol.* **2017**, *2017*, 1–21, doi:10.1002/cptx.19.

258. Takumi, T.; Tamada, K.; Hatanaka, F.; Nakai, N.; Bolton, P.F. Behavioral neuroscience of autism. *Neurosci. Biobehav. Rev.* **2020**, *110*, 60–76, doi:10.1016/j.neubiorev.2019.04.012.

259. Meyza, K.Z.; Defensor, E.B.; Jensen, A.L.; Corley, M.J.; Pearson, B.L.; Pobbe, R.L.H.; Bolivar, V.J.; Blanchard, D.C.; Blanchard, R.J. The BTBR T+tf/J mouse model for autism spectrum disorders-in search of biomarkers. *Behav. Brain Res.* **2013**, *251*, 25–34, doi:10.1016/j.bbr.2012.07.021.

260. De La Torre-Ubieta, L.; Won, H.; Stein, J.L.; Geschwind, D.H. Advancing the understanding of autism disease mechanisms through genetics. *Nat. Med.* **2016**, *22*, 345–361, doi:10.1038/nm.4071.

261. Ergaz, Z.; Weinstein-Fudim, L.; Ornoy, A. Genetic and non-genetic animal models for autism spectrum disorders (ASD). *Reprod. Toxicol.* **2016**, *64*, 116–140, doi:10.1016/j.reprotox.2016.04.024.

262. Norton, W. Towards developmental models of psychiatric disorders in zebrafish. *Front. Neural Circuits* **2013**, *7*, 1–12, doi:10.3389/fncir.2013.00079.

263. Stewart, A.M.; Ullmann, J.F.P.; Norton, W.H.J.; Brennan, C.H.; Parker, M.O.; Gerlai, R.; Kalueff, A. V Molecular psychiatry of zebrafish. *Mol. Psychiatry* **2015**, *20*, 2–17, doi:10.1038/mp.2014.128.Molecular.

264. Fontana, B.D.; Mezzomo, N.J.; Kalueff, A. V.; Rosemberg, D.B. The developing utility of zebrafish models of neurological and neuropsychiatric disorders: A critical review. *Exp. Neurol.* **2018**, *299*, 157–171, doi:10.1016/j.expneurol.2017.10.004.

265. Kalueff, A. V; Stewart, A.M.; Gerlai, R. Zebrafish as an emerging model for studying complex brain disorders. **2014**, *35*, 63–75, doi:10.1016/j.tips.2013.12.002.Zebrafish.

266. Kalueff, A. V.; Echevarria, D.J.; Stewart, A.M. Gaining translational momentum: More zebrafish models for

neuroscience research. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2014**, *55*, 1–6, doi:10.1016/j.pnpbp.2014.01.022.

267. Howe, K.; Clark, M.D.; Torroja, C.F.; Torrance, J.; Berthelot, C.; Muffato, M.; Collins, J.E.; Humphray, S.; McLaren, K.; Matthews, L.; et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **2013**, *496*, 498–503, doi:10.1038/nature12111.

268. Tropepe, V.; Sive, H.L. Can zebrafish be used as a model to study the neurodevelopmental causes of autism? *Genes, Brain Behav.* **2003**, *2*, 268–281, doi:10.1034/j.1601-183X.2003.00038.x.

269. Vaz, R.; Hofmeister, W.; Lindstrand, A. Zebrafish models of neurodevelopmental disorders: Limitations and benefits of current tools and techniques. *Int. J. Mol. Sci.* **2019**, *20*, doi:10.3390/ijms20061296.

270. Kozol, R.A.; Abrams, A.J.; James, D.M.; Buglo, E.; Yan, Q.; Dallman, J.E. Function over form: Modeling groups of inherited neurological conditions in zebrafish. *Front. Mol. Neurosci.* **2016**, *9*, 1–15, doi:10.3389/fnmol.2016.00055.

271. Meshalkina, D.A.; N. Kizlyk, M.; V. Kysil, E.; Collier, A.D.; Echevarria, D.J.; Abreu, M.S.; Barcellos, L.J.G.; Song, C.; Warnick, J.E.; Kyzar, E.J.; et al. Zebrafish models of autism spectrum disorder. *Exp. Neurol.* **2018**, *299*, 207–216, doi:10.1016/j.expneurol.2017.02.004.

272. Sakai, C.; Ijaz, S.; Hoffman, E.J. Zebrafish Models of Neurodevelopmental Disorders: Past, Present, and Future. *Front. Mol. Neurosci.* **2018**, *11*, doi:10.3389/fnmol.2018.00294.

273. Rea, V.; Van Raay, T.J. Using Zebrafish to Model Autism Spectrum Disorder: A Comparison of ASD Risk Genes Between Zebrafish and Their Mammalian Counterparts. *Front. Mol. Neurosci.* **2020**, *13*, 1–19, doi:10.3389/fnmol.2020.575575.

274. Hillman, E.M.C. Coupling mechanism and significance of the BOLD signal: A status report. *Annu. Rev. Neurosci.* 2014, *37*, 161–181.

275. Li, X.; Zhang, K.; He, X.; Zhou, J.; Jin, C. Structural , Functional , and Molecular Imaging of Autism Spectrum Disorder. *Neurosci. Bull.* **2021**, doi:10.1007/s12264-021-00673-0.

276. Wolff, J.J.; Gu, H.; Gerig, G.; Elison, J.T.; Styner, M.; Gouttard, S.; Botteron, K.N.; Dager, S.R.; Dawson, G.; Estes, A.M.; et al. Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *Am. J. Psychiatry* **2012**, *169*, 589–600, doi:10.1176/appi.ajp.2011.11091447.

277. Wolff, J.J.; Swanson, M.R.; Elison, J.T.; Gerig, G.; Pruett, J.R.; Styner, M.A.; Vachet, C.; Botteron, K.N.; Dager, S.R.; Estes, A.M.; et al. Neural circuitry at age 6 months associated with later repetitive behavior and sensory responsiveness in autism. *Mol. Autism* **2017**, *8*, doi:10.1186/s13229-017-0126-z.

278. Smith, E.; Thurm, A.; Greenstein, D.; Farmer, C.; Swedo, S.; Giedd, J.; Raznahan, A. Cortical thickness change in autism during early childhood. *Hum. Brain Mapp.* **2016**, *37*, 2616–2629, doi:10.1002/hbm.23195.

279. Deramus, T.P.; Kana, R.K. Anatomical likelihood estimation meta-analysis of grey and white matter anomalies in autism spectrum disorders. *NeuroImage Clin.* **2015**, *7*, 525–536, doi:10.1016/j.nicl.2014.11.004.

280. Liu, J.; Yao, L.; Zhang, W.; Xiao, Y.; Liu, L.; Gao, X.; Shah, C.; Li, S.; Tao, B.; Gong, Q.; et al. Gray matter abnormalities in pediatric autism spectrum disorder: a meta-analysis with signed differential mapping. *Eur. Child Adolesc. Psychiatry* **2017**, *26*, 933–945, doi:10.1007/s00787-017-0964-4.

281. Schumann, C.M.; Bloss, C.S.; Barnes, C.C.; Wideman, G.M.; Carper, R.A.; Akshoomoff, N.; Pierce, K.; Hagler, D.; Schork, N.; Lord, C.; et al. Longitudinal magnetic resonance imaging study of cortical development through early

childhood in autism. *J. Neurosci.* **2010**, *30*, 4419–4427, doi:10.1523/JNEUROSCI.5714-09.2010.

282. Van Rooij, D.; Anagnostou, E.; Arango, C.; Auzias, G.; Behrmann, M.; Busatto, G.F.; Calderoni, S.; Daly, E.; Deruelle, C.; Di Martino, A.; et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD working group. *Am. J. Psychiatry* **2018**, *175*, 359–369, doi:10.1176/appi.ajp.2017.17010100.

283. Robertson, C.E.; Baron-Cohen, S. Sensory perception in autism. *Nat. Rev. Neurosci.* **2017**, *18*, 671–684, doi:10.1038/nrn.2017.112.

284. Herringshaw, A.J.; Ammons, C.J.; DeRamus, T.P.; Kana, R.K. Hemispheric differences in language processing in autism spectrum disorders: A meta-analysis of neuroimaging studies. *Autism Res.* **2016**, *9*, 1046–1057, doi:10.1002/aur.1599.

285. Dichter, G.S. Functional magnetic resonance imaging of autism spectrum disorders. *Dialogues Clin. Neurosci.* **2012**, *14*, 319–351, doi:10.31887/dcns.2012.14.3/gdichter.

286. Di Martino, A.; Yan, C.G.; Li, Q.; Denio, E.; Castellanos, F.X.; Alaerts, K.; Anderson, J.S.; Assaf, M.; Bookheimer, S.Y.; Dapretto, M.; et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* **2014**, *19*, 659–667, doi:10.1038/mp.2013.78.

287. Doyle-Thomas, K.A.R.; Lee, W.; Foster, N.E.V.; Tryfon, A.; Ouimet, T.; Hyde, K.L.; Evans, A.C.; Lewis, J.; Zwaigenbaum, L.; Anagnostou, E. Atypical functional brain connectivity during rest in autism spectrum disorders. *Ann. Neurol.* **2015**, *77*, 866–876, doi:10.1002/ana.24391.

288. Supekar, K.; Uddin, L.Q.; Khouzam, A.; Phillips, J.; Gaillard, W.D.; Kenworthy, L.E.; Yerys, B.E.; Vaidya, C.J.; Menon, V. Brain Hyperconnectivity in Children with Autism and its Links to Social Deficits. *Cell Rep.* **2013**, *5*, 738–747, doi:10.1016/j.celrep.2013.10.001.

289. Jann, K.; Hernandez, L.M.; Beck-Pancer, D.; Mccarron, R.; Smith, R.X.; Dapretto, M.; Wang, D.J.J. Altered resting perfusion and functional connectivity of default mode network in youth with autism spectrum disorder. *Brain Behav.* **2015**, *5*, doi:10.1002/brb3.358.

290. Peterson, B.S.; Zargarian, A.; Peterson, J.B.; Goh, S.; Sawardekar, S.; Williams, S.C.R.; Lythgoe, D.J.; Zelaya, F.O.; Bansal, R. Hyperperfusion of Frontal White and Subcortical Gray Matter in Autism Spectrum Disorder. *Biol. Psychiatry* **2019**, *85*, 584–595, doi:10.1016/j.biopsych.2018.11.026.

291. Mitelman, S.A.; Bralet, M.C.; Mehmet Haznedar, M.; Hollander, E.; Shihabuddin, L.; Hazlett, E.A.; Buchsbaum, M.S. Positron emission tomography assessment of cerebral glucose metabolic rates in autism spectrum disorder and schizophrenia. *Brain Imaging Behav.* **2018**, *12*, 532–546, doi:10.1007/s11682-017-9721-z.

292. Mitelman, S.A.; Buchsbaum, M.S.; Young, D.S.; Haznedar, M.M.; Hollander, E.; Shihabuddin, L.; Hazlett, E.A.; Bralet, M.C. Increased white matter metabolic rates in autism spectrum disorder and schizophrenia. *Brain Imaging Behav.* **2018**, *12*, 1290–1305, doi:10.1007/s11682-017-9785-9.

293. Oberman, L.M.; Hubbard, E.M.; McCleery, J.P.; Altschuler, E.L.; Ramachandran, V.S.; Pineda, J.A. EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cogn. Brain Res.* **2005**, *24*, 190–198, doi:10.1016/j.cogbrainres.2005.01.014.

294. Southgate, V.; Hamilton, A.F. d. C. Unbroken mirrors: challenging a theory of Autism. *Trends Cogn. Sci.* **2008**, *12*, 225–229, doi:10.1016/j.tics.2008.03.005.

295. Fan, Y.T.; Decety, J.; Yang, C.Y.; Liu, J.L.; Cheng, Y. Unbroken mirror neurons in autism spectrum disorders. *J.*

*Child Psychol. Psychiatry Allied Discip.* **2010**, *51*, 981–988, doi:10.1111/j.1469-7610.2010.02269.x.

296. Bernier, R.; Aaronson, B.; McPartland, J. The role of imitation in the observed heterogeneity in EEG mu rhythm in autism and typical development. *Brain Cogn.* **2013**, *82*, 69–75, doi:10.1016/j.bandc.2013.02.008.

297. Raymaekers, R.; Wiersema, J.R.; Roeyers, H. EEG study of the mirror neuron system in children with high functioning autism. *Brain Res.* **2009**, *1304*, 113–121, doi:10.1016/j.brainres.2009.09.068.

298. Dumas, G.; Soussignan, R.; Hugueville, L.; Martinerie, J.; Nadel, J. Revisiting mu suppression in autism spectrum disorder. *Brain Res.* **2014**, *1585*, 108–119, doi:10.1016/j.brainres.2014.08.035.

299. Black, M.H.; Chen, N.T.M.; Iyer, K.K.; Lipp, O. V.; Bölte, S.; Falkmer, M.; Tan, T.; Girdler, S. Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neurosci. Biobehav. Rev.* 2017, *80*, 488–515.

300. Marco, E.J.; Hinkley, L.B.N.; Hill, S.S.; Nagarajan, S.S. Sensory processing in autism: A review of neurophysiologic findings. *Pediatr. Res.* **2011**, *69*, doi:10.1203/PDR.0b013e3182130c54.

301. Tierney, A.L.; Gabard-Durnam, L.; Vogel-Farley, V.; Tager-Flusberg, H.; Nelson, C.A. Developmental Trajectories of Resting EEG Power: An Endophenotype of Autism Spectrum Disorder. *PLoS One* **2012**, *7*, e39127, doi:10.1371/journal.pone.0039127.

302. Gabard-Durnam, L.; Tierney, A.L.; Vogel-Farley, V.; Tager-Flusberg, H.; Nelson, C.A. Alpha asymmetry in infants at risk for autism spectrum disorders. *J. Autism Dev. Disord.* **2015**, *45*, 473–480, doi:10.1007/s10803-013-1926-4.

303. Dana, H.; Mohar, B.; Sun, Y.; Narayan, S.; Gordus, A.; Hasseman, J.P.; Tsegaye, G.; Holt, G.T.; Hu, A.; Walpita, D.; et al. Sensitive red protein calcium indicators for imaging neural activity. *Elife* **2016**, *5*, 1–24, doi:10.7554/eLife.12727.

304. Chen, T.; Wardill, T.J.; Sun, Y.; Pulver, S.R.; Renninger, S.L.; Baohan, A.; Schreiter, E.R.; Kerr, R.A.; Orger, M.B.; Jayaraman, V.; et al. Ultra-sensitive fluorescent proteins for imaging neuronal activity. *Nature* **2013**, *499*, 295–300.

305. Zhang, Z.; Cong, L.; Bai, L.; Wang, K. Light-field microscopy for fast volumetric brain imaging. *J. Neurosci. Methods* **2021**, *352*, 109083, doi:10.1016/j.jneumeth.2021.109083.

306. Ahrens, M.B.; Orger, M.B.; Robson, D.N.; Li, J.M.; Keller, P.J. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods* **2013**, *10*, 413–420, doi:10.1038/nmeth.2434.

307. Ahrens, M.B.; Huang, K.H.; Narayan, S.; Mensh, B.D.; Engert, F. Two-photon calcium imaging during fictive navigation in virtual environments. *Front. Neural Circuits* **2013**, *7*, doi:10.3389/fncir.2013.00104.

308. Vladimirov, N.; Mu, Y.; Kawashima, T.; Bennett, D. V.; Yang, C.T.; Looger, L.L.; Keller, P.J.; Freeman, J.; Ahrens, M.B. Light-sheet functional imaging in fictively behaving zebrafish. *Nat. Methods* **2014**, *11*, 883–884, doi:10.1038/nmeth.3040.

309. Keller, P.J.; Ahrens, M.B. Visualizing whole-brain activity and development at the single-cell level using light-sheet microscopy. *Neuron* **2015**, *85*, 462–483, doi:10.1016/j.neuron.2014.12.039.

310. Vanwalleghem, G.C.; Ahrens, M.B.; Scott, E.K. Integrative whole-brain neuroscience in larval zebrafish. *Curr. Opin. Neurobiol.* **2018**, *50*, 136–145, doi:10.1016/j.conb.2018.02.004.

311. Lin, Q.; Manley, J.; Helmreich, M.; Schlumm, F.; Li, J.M.; Robson, D.N.; Engert, F.; Schier, A.; Nöbauer, T.; Vaziri,

A. Cerebellar Neurodynamics Predict Decision Timing and Outcome on the Single-Trial Level. *Cell* **2020**, *180*, 536-551.e17, doi:10.1016/j.cell.2019.12.018.

312. Light, S.E.W.; Jontes, J.D. Multiplane calcium imaging reveals disrupted development of network topology in zebrafish pcdh19 mutants. *eNeuro* **2019**, *6*, 1–15, doi:10.1523/ENEURO.0420-18.2019.

313. Reichmann, F.; Rimmer, N.; Tilley, C.A.; Dalla Vecchia, E.; Pinion, J.; Al Oustah, A.; Carreño Gutiérrez, H.; Young, A.M.J.; McDearmid, J.R.; Winter, M.J.; et al. The zebrafish histamine H3 receptor modulates aggression, neural activity and forebrain functional connectivity. *Acta Physiol.* **2020**, *230*, 1–17, doi:10.1111/apha.13543.

314. World Health Organization *The ICD-10 classification of mental and behavioural disorders*; World Health Organization, 1993; ISBN 9789241544559.

315. Bussu, G.; Jones, E.J.H.; Charman, T.; Johnson, M.H.; Buitelaar, J.K.; Blasi, A.; Baron-Cohen, S.; Bedford, R.; Bolton, P.; Chandler, S.; et al. Latent trajectories of adaptive behaviour in infants at high and low familial risk for autism spectrum disorder. *Mol. Autism* **2019**, *10*, 1–12, doi:10.1186/s13229-019-0264-6.

316. Zerbi, V.; Ielacqua, G.D.; Markicevic, M.; Haberl, M.G.; Ellisman, M.H.; Bhaskaran, A.A.; Frick, A.; Rudin, M.; Wenderoth, N. Dysfunctional autism risk genes cause circuit-specific connectivity deficits with distinct developmental trajectories. *Cereb. Cortex* **2018**, *28*, 2495–2506, doi:10.1093/cercor/bhy046.

317. World Health Organization *ICD-11 for Mortality and Morbidity Statistics*; World Health Organiztion, 2018;

318. Kozak, M.J.; Cuthbert, B.N. The NIMH Research Domain Criteria Initiative: Background, Issues, and Pragmatics. *Psychophysiology* **2016**, *53*, 286–297, doi:10.1111/psyp.12518.

319. Sparrow, S.S.; Cicchetti, D. V; Balla, D.A. *Vineland Adaptive Behavior Scales*; Pearson.; 2005;

320. Nsengimana, J.; Bishop, T. Design Considerations for Genetic Linkage and Association Studies. *Stat. Hum. Genet. Methods Protoc. Methods Mol. Biol.* **2017**, *1666*, 257–281, doi:10.4324/9781351218986-13.

321. Balding, D.J. A tutorial on statistical methods for population association studies. *Nat. Rev. Genet.* **2006**, *7*, 781–791, doi:10.1038/nrg1916.

322. Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **2019**, *20*, 467–484, doi:10.1038/s41576-019-0127-1.

323. McCarthy, M.I.; Abecasis, G.R.; Cardon, L.R.; Goldstein, D.B.; Little, J.; Ioannidis, J.P.A.; Hirschhorn, J.N. Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat. Rev. Genet.* **2008**, *9*, 356–369, doi:10.1038/nrg2344.

324. Narita, A.; Nagai, M.; Mizuno, S.; Ogishima, S.; Tamiya, G.; Ueki, M.; Sakurai, R.; Makino, S.; Obara, T.; Ishikuro, M.; et al. Clustering by phenotype and genome-wide association study in autism. *Transl. Psychiatry* **2020**, *10*, doi:10.1038/s41398-020-00951-x.

325. Mattheisen, M.; Grove, J.; Als, T.D.; Martin, J.; Voloudakis, G.; Meier, S.; Demontis, D.; Bendl, J.; Walters, R.; Carey, C.E.; et al. Identification of shared and differentiating genetic risk for autism spectrum disorder , attention deficit hyperactivity disorder and case subgroups. *medRxiv* **2021**, doi:https://doi.org/10.1101/2021.05.20.21257484.

326. Hoggart, C.J.; Clark, T.G.; De Iorio, M.; Whittaker, J.C.; Balding, D.J. Genome-wide significance for dense SNP and resequencing data. *Genet. Epidemiol.* **2008**, *32*, 179–185, doi:10.1002/gepi.20292.

327. Ripke, S.; Neale, B.M.; Corvin, A.; Walters, J.T.R.; Farh, K.H.; Holmans, P.A.; Lee, P.; Bulik-Sullivan, B.; Collier,

D.A.; Huang, H.; et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **2014**, *511*, 421–427, doi:10.1038/nature13595.

328. de Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **2015**, *11*, 4, doi:10.1371/journal.pcbi.1004219.

329. Kido, M.; Nakamura, Y.; Nemoto, K.; Takahashi, T.; Aleksic, B.; Furuichi, A.; Nakamura, Y.; Ikeda, M.; Noguchi, K.; Kaibuchi, K.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and brain morphology in schizophrenia: A voxel-based morphometric study. *PLoS One* **2014**, *9*, E103571, doi:10.1371/journal.pone.0103571.

330. Takahashi, T.; Nakamura, Y.; Nakamura, Y.; Aleksic, B.; Takayanagi, Y.; Furuichi, A.; Kido, M.; Nakamura, M.; Sasabayashi, D.; Ikeda, M.; et al. The polymorphism of YWHAE, a gene encoding 14-3-3epsilon, and orbitofrontal sulcogyral pattern in patients with schizophrenia and healthy subjects. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2014**, *51*, 166–171, doi:10.1016/j.pnpbp.2014.02.005.

331. Wise, A.L.; Gyi, L.; Manolio, T.A. EXclusion: Toward integrating the X chromosome in genome-wide association analyses. *Am. J. Hum. Genet.* **2013**, *92*, 643–647, doi:10.1016/j.ajhg.2013.03.017.

332. Bai, D.; Yip, B.H.K.; Windham, G.C.; Sourander, A.; Francis, R.; Yoffe, R.; Glasson, E.; Mahjani, B.; Suominen, A.; Leonard, H.; et al. Association of Genetic and Environmental Factors with Autism in a 5-Country Cohort. *JAMA Psychiatry* **2019**, *76*, 1035–1043, doi:10.1001/jamapsychiatry.2019.1411.

333. An, J.Y.; Lin, K.; Zhu, L.; Werling, D.M.; Dong, S.; Brand, H.; Wang, H.Z.; Zhao, X.; Schwartz, G.B.; Collins, R.L.; et al. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. *Science (80-. ).* **2018**, *362*, doi:10.1126/science.aat6576.

334. Sonuga-Barke, E.J.S.; Lasky-Su, J.; Neale, B.M.; Oades, R.; Chen, W.; Franke, B.; Buitelaar, J.; Banaschewski, T.; Ebstein, R.; Gill, M.; et al. Does parental expressed emotion moderate genetic effects in ADHD? An exploration using a genome wide association scan. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2008**, *147*, 1359–1368, doi:10.1002/ajmg.b.30860.

335. Richter, J.; Hamm, A.O.; Pan-Farr, C.A.; Gerlach, A.L.; Gloster, A.T.; Wittchen, H.U.; Lang, T.; Alpers, G.W.; Helbig-Lang, S.; Deckert, J.; et al. Dynamics of defensive reactivity in patients with panic disorder and agoraphobia: Implications for the etiology of panic disorder. *Biol. Psychiatry* **2012**, *72*, 512–520, doi:10.1016/j.biopsych.2012.03.035.

336. Vogt, B.A. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* 2005, *6*, 533–544.

337. Laird, A.R.; Fox, P.M.; Eickhoff, S.B.; Turner, J.A.; Ray, K.L.; Mckay, D.R.; Glahn, D.C.; Beckmann, C.F.; Smith, S.M.; Fox, P.T. *Behavioral Interpretations of Intrinsic Connectivity Networks*; 2011;

338. Fullana, M.A.; Harrison, B.J.; Soriano-Mas, C.; Vervliet, B.; Cardoner, N.; Àvila-Parcet, A.; Radua, J. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* **2016**, *21*, 500–508, doi:10.1038/mp.2015.88.

339. Etkin, A.; Egner, T.; Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **2011**, *15*, 85–93, doi:10.1016/j.tics.2010.11.004.

340. Milad, M.R.; Quirk, G.J. Fear extinction as a model for translational neuroscience: Ten years of progress. *Annu. Rev. Psychol.* 2012, *63*, 129–151.

REFERENCES

341. Milad, M.R.; Quirk, G.J.; Pitman, R.K.; Orr, S.P.; Fischl, B.; Rauch, S.L. A Role for the Human Dorsal Anterior Cingulate Cortex in Fear Expression. *Biol. Psychiatry* **2007**, *62*, 1191–1194, doi:10.1016/j.biopsych.2007.04.032.

342. Sehlmeyer, C.; Schöning, S.; Zwitserlood, P.; Pfleiderer, B.; Kircher, T.; Arolt, V.; Konrad, C. Human fear conditioning and extinction in neuroimaging: A systematic review. *PLoS One* 2009, *4*, 5865.

343. Badre, D. Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn. Sci.* 2008, *12*, 193–200.

344. Swick, D.; Ashley, V.; Turken, A.U. Left inferior frontal gyrus is critical for response inhibition. **2008**, doi:10.1186/1471-2202-9-102.

345. Yücel, M.; Wood, S.J.; Fornito, A.; Riffkin, J.; Velakoulis, D.; Pantelis, C. *Anterior cingulate dysfunction: Implications for psychiatric disorders?*; 2003; Vol. 28;.

346. Perach-Barzilay, N.; Tauber, A.; Klein, E.; Chistyakov, A.; Ne'eman, R.; Shamay-Tsoory, S.G. Asymmetry in the dorsolateral prefrontal cortex and aggressive behavior: A continuous theta-burst magnetic stimulation study. *Soc. Neurosci.* **2013**, *8*, 178–188, doi:10.1080/17470919.2012.720602.

347. Sanders, S.J.; Neale, B.M.; Huang, H.; Werling, D.M.; An, J.Y.; Dong, S.; Abecasis, G.; Arguello, P.A.; Blangero, J.; Boehnke, M.; et al. Whole genome sequencing in psychiatric disorders: The WGSPD consortium. *Nat. Neurosci.* **2017**, *20*, 1661–1668, doi:10.1038/s41593-017-0017-9.

348. Satterstrom, F.K.; Walters, R.K.; Singh, T.; Wigdor, E.M.; Lescai, F.; Demontis, D.; Kosmicki, J.A.; Grove, J.; Stevens, C.; Bybjerg-Grauholm, J.; et al. Autism spectrum disorder and attention deficit hyperactivity disorder have a similar burden of rare protein-truncating variants. *Nat. Neurosci.* **2019**, *22*, 1961–1965, doi:10.1038/s41593-019-0527-8.

349. Cornell, B.; Wachi, T.; Zhukarev, V.; Toyo-oka, K. Regulation of neuronal morphogenesis by 14-3-3epsilon ( Ywhae ) via the microtubule binding protein, doublecortin . *Hum. Mol. Genet.* **2016**, *25*, 4405–4418, doi:10.1093/hmg/ddw270.

350. Bruno, D.L.; Anderlid, B.M.; Lindstrand, A.; Van Ravenswaaij-Arts, C.; Ganesamoorthy, D.; Lundin, J.; Martin, C.L.; Douglas, J.; Nowak, C.; Adam, M.P.; et al. Further molecular and clinical delineation of co-locating 17p13.3 microdeletions and microduplications that show distinctive phenotypes. *J. Med. Genet.* **2010**, *47*, 299–311, doi:10.1136/jmg.2009.069906.

351. Capra, V.; Mirabelli-Badenier, M.; Stagnaro, M.; Rossi, A.; Tassano, E.; Gimelli, S.; Gimelli, G. Identification of a rare 17p13.3 duplication including the BHLHA9 and YWHAE genes in a family with developmental delay and behavioural problems. *BMC Med. Genet.* **2012**, *13*, 93, doi:10.1186/1471-2350-13-93.

352. Nagamani, S.C.S.; Zhang, F.; Shchelochkov, O.A.; Bi, W.; Ou, Z.; Scaglia, F.; Probst, F.J.; Shinawi, M.; Eng, C.; Hunter, J. V.; et al. Microdeletions including YWHAE in the Miller-Dieker syndrome region on chromosome 17p13.3 result in facial dysmorphisms, growth restriction, and cognitive impairment. *J. Med. Genet.* **2009**, *46*, 825–833, doi:10.1136/jmg.2009.067637.

353. Blazejewski, S.M.; Bennison, S.A.; Smith, T.H.; Toyo-oka, K. Neurodevelopmental genetic diseases associated with microdeletions and microduplications of chromosome 17p13.3. *Front. Genet.* **2018**, *9*, 9, doi:10.3389/fgene.2018.00080.

354. Curry, C.J.; Rosenfeld, J.A.; Grant, E.; Gripp, K.W.; Anderson, C.; Aylsworth, A.S.; Saad, T. Ben; Chizhikov, V. V.; Dybose, G.; Fagerberg, C.; et al. The duplication 17p13.3 phenotype: Analysis of 21 families delineates

developmental, behavioral and brain abnormalities, and rare variant phenotypes. *Am. J. Med. Genet. Part A* **2013**, *161*, 1833–1852, doi:10.1002/ajmg.a.35996.

355. Foote, M.; Zhou, Y. 14-3-3 Proteins in Neurological Disorders. *Int. J. Biochem. Mol. Biol.* **2012**, *3*, 152–164.

356. Fromer, M.; Pocklington, A.J.; Kavanagh, D.H.; Williams, H.J.; Dwyer, S.; Gormley, P.; Georgieva, L.; Rees, E.; Palta, P.; Ruderfer, D.M.; et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature* **2014**, *506*, 179–184, doi:10.1038/nature12929.

357. Purcell, S.M.; Moran, J.L.; Fromer, M.; Ruderfer, D.; Solovieff, N.; Roussos, P.; O'Dushlaine, C.; Chambert, K.; Bergen, S.E.; Kähler, A.; et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **2014**, *506*, 185–190, doi:10.1038/nature12975.

358. Du, X.; An, Y.; Yu, L.; Liu, R.; Qin, Y.; Guo, X.; Sun, D.; Zhou, S.; Wu, B.; Jiang, Y. hui; et al. A genomic copy number variant analysis implicates the MBD5 and HNRNPU genes in Chinese children with infantile spasms and expands the clinical spectrum of 2q23.1 deletion. *BMC Med. Genet.* **2014**, *15*, 1–12, doi:10.1186/1471-2350-15-62.

359. Halskau, Ø.J.; Ying, M.; Baumann, A.; Kleppe, R.; Rodriguez-Larrea, D.; Almås, B.; Haavik, J.; Martinez, A. Three-way interaction between 14-3-3 proteins, the N-terminal region of tyrosine hydroxylase, and negatively charged membranes. *J. Biol. Chem.* **2009**, *284*, 32758–32769, doi:10.1074/jbc.M109.027706.

360. Ghorbani, S.; Fossbakk, A.; Jorge-Finnigan, A.; Flydal, M.I.; Haavik, J.; Kleppe, R. Regulation of tyrosine hydroxylase is preserved across different homo- and heterodimeric 14-3-3 proteins. *Amino Acids* **2016**, *48*, 1221–1229, doi:10.1007/s00726-015-2157-0.

361. Kesby, J.P.; Eyles, D.W.; McGrath, J.J.; Scott, J.G. Dopamine, psychosis and schizophrenia: The widening gap between basic and clinical neuroscience. *Transl. Psychiatry* 2018, *8*, 30.

362. Kim, D.; Won, J.; Shin, D.W.; Kang, J.; Kim, Y.J.; Choi, S.Y.; Hwang, M.K.; Jeong, B.W.; Kim, G.S.; Joe, C.O.; et al. Regulation of Dyrk1A kinase activity by 14-3-3. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 499–504, doi:10.1016/j.bbrc.2004.08.102.

363. Benzinger, A.; Muster, N.; Koch, H.B.; Yates, J.R.; Hermeking, H. Targeted proteomic analysis of 14-3-3σ, a p53 effector commonly silenced in cancer. *Mol. Cell. Proteomics* **2005**, *4*, 785–795, doi:10.1074/mcp.M500021-MCP200.

364. Alvarez, M.; Altafaj, X.; Aranda, S.; de la Luna, S. DYRK1A Autophosphorylation on Serine Residue 520 Modulates Its Kinase Activity via 14-3-3 Binding. *Mol. Biol. Cell* **2007**, *18*, 1167–1178, doi:10.1091/mbc.E06.

365. Jin, J.; Smith, F.D.; Stark, C.; Wells, lark D.; Fawcett, ames P.; Kulkarni, S.; Metalnikov, A.; O'Donnell, P.; Taylor, P.; Taylor, L.; et al. Proteomic, Functional, and Domain-Based Analysis of In Vivo 14-3-3 Binding Proteins Involved in Cytoskeletal Regulation and Cellular Organization. *Curr. Biol.* **2004**, *14*, 1436–1450, doi:10.1016/j.cub.2004.07.051.

366. Arbones, M.L.; Thomazeau, A.; Nakano-Kobayashi, A.; Hagiwara, M.; Delabar, J.M. DYRK1A and cognition: A lifelong relationship. *Pharmacol. Ther.* **2019**, *194*, 199–221, doi:10.1016/j.pharmthera.2018.09.010.

367. Dang, T.; Duan, W.Y.; Yu, B.; Tong, D.L.; Cheng, C.; Zhang, Y.F.; Wu, W.; Ye, K.; Zhang, W.X.; Wu, M.; et al. Autism-associated Dyrk1a truncation mutants impair neuronal dendritic and spine growth and interfere with postnatal cortical development. *Mol. Psychiatry* **2018**, *23*, 747–758, doi:10.1038/mp.2016.253.

368. Egervari, G.; Kozlenkov, A.; Dracheva, S.; Hurd, Y.L. Molecular windows into the human brain for psychiatric disorders. *Mol. Psychiatry* **2019**, *24*, 653–673, doi:10.1038/s41380-018-0125-2.

369. Hernandez, L.M.; Kim, M.; Hoftman, G.D.; Haney, J.R.; de la Torre-Ubieta, L.; Pasaniuc, B.; Gandal, M.J. Transcriptomic Insight Into the Polygenic Mechanisms Underlying Psychiatric Disorders. *Biol. Psychiatry* **2021**, *89*, 54–64, doi:10.1016/j.biopsych.2020.06.005.

370. Cáceres, M.; Lachuer, J.; Zapala, M.A.; Redmond, J.C.; Kudo, L.; Geschwind, D.H.; Lockhart, D.J.; Preuss, T.M.; Barlow, C. Elevated gene expression levels distinguish human from non-human primate brains. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13030–13035, doi:10.1073/pnas.2135499100.

371. Konopka, G.; Friedrich, T.; Davis-Turak, J.; Winden, K.; Oldham, M.C.; Gao, F.; Chen, L.; Wang, G.Z.; Luo, R.; Preuss, T.M.; et al. Human-Specific Transcriptional Networks in the Brain. *Neuron* **2012**, *75*, 601–617, doi:10.1016/j.neuron.2012.05.034.

372. Skefos, J.; Cummings, C.; Enzer, K.; Holiday, J.; Weed, K.; Levy, E.; Yuce, T.; Kemper, T.; Bauman, M. Regional Alterations in Purkinje Cell Density in Patients with Autism. *PLoS One* **2014**, *9*, e81255, doi:10.1371/journal.pone.0081255.

373. Chugani, D.C. Role of altered brain serotonin mechanisms in autism. *Mol. Psychiatry* **2002**, *7*, 16–17, doi:10.1038/sj.mp.4001167.

374. Chugani, D.C.; Muzik, O.; Behen, M.; Rothermel, R.; Janisse, J.J.; Lee, J.; Chugani, H.T. Developmental changes in brain serotonin synthesis capacity in autistic and nonautistic children. *Ann. Neurol.* **1999**, *45*, 287–295, doi:10.1002/1531-8249(199903)45:3<287::AID-ANA3>3.0.CO;2-9.

375. Dichter, G.S.; Felder, J.N.; Green, S.R.; Rittenberg, A.M.; Sasson, N.J.; Bodfish, J.W. Reward circuitry function in autism spectrum disorders. *Soc. Cogn. Affect. Neurosci.* **2012**, *7*, 160–172, doi:10.1093/scan/nsq095.

376. Neuhaus, E.; Beauchaine, T.P.; Bernier, R. Neurobiological correlates of social functioning in autism. *Clin. Psychol. Rev.* **2010**, *30*, 733–748, doi:10.1016/j.cpr.2010.05.007.

377. Scott-Van Zeeland, A.A.; Dapretto, M.; Ghahremani, D.G.; Poldrack, R.A.; Bookheimer, S.Y. Reward processing in autism. *Autism Res.* **2010**, *3*, 53–67, doi:10.1002/aur.122.

378. Vuong, C.K.; Wei, W.; Lee, J.A.; Lin, C.H.; Damianov, A.; de la Torre-Ubieta, L.; Halabi, R.; Otis, K.O.; Martin, K.C.; O'Dell, T.J.; et al. Rbfox1 Regulates Synaptic Transmission through the Inhibitory Neuron-Specific vSNARE Vamp1. *Neuron* **2018**, *98*, 127-141.e7, doi:10.1016/j.neuron.2018.03.008.

379. Hamada, N.; Ito, H.; Iwamoto, I.; Morishita, R.; Tabata, H.; Nagata, K.I. Role of the cytoplasmic isoform of RBFOX1/A2BP1 in establishing the architecture of the developing cerebral cortex. *Mol. Autism* **2015**, *6*, 1–13, doi:10.1186/s13229-015-0049-5.

380. Teles, M.C.; Cardoso, S.D.; Oliveira, R.F. Social plasticity relies on different neuroplasticity mechanisms across the brain social decision-making network in zebrafish. *Front. Behav. Neurosci.* **2016**, *10*, doi:10.3389/fnbeh.2016.00016.

381. Mueller, T. What is the thalamus in zebrafish? *Front. Neurosci.* **2012**, *6*, 1–14, doi:10.3389/fnins.2012.00064.

382. Qin, C.; Li, J.; Tang, K. The paraventricular nucleus of the hypothalamus: Development, function, and human diseases. *Endocrinology* **2018**, *159*, 3458–3472, doi:10.1210/en.2018-00453.

383. Stednitz, S.J.; McDermott, E.M.; Ncube, D.; Tallafuss, A.; Eisen, J.S.; Washbourne, P. Forebrain Control of Behaviorally Driven Social Orienting in Zebrafish. *Curr. Biol.* **2018**, *28*, 2445-2451.e3, doi:10.1016/j.cub.2018.06.016.

384. Fombonne, E.; Rogé, B.; Claverie, J.; Courty, S.; Frémolle, J. Microcephaly and macrocephaly in autism. *J. Autism*

*Dev. Disord.* **1999**, *29*, 113–119, doi:10.1023/A:1023036509476.

385. Wright, I.C.; Rabe-Hesketh, S.; Woodruff, P.W.R.; David, A.S.; Murray, R.M.; Bullmore, E.T. Meta-analysis of regional brain volumes in schizophrenia. *Am. J. Psychiatry* **2000**, *157*, 16–25, doi:10.1176/ajp.157.1.16.

386. Turner, A.H.; Greenspan, K.S.; van Erp, T.G.M. Pallidum and lateral ventricle volume enlargement in autism spectrum disorder. *Psychiatry Res. - Neuroimaging* **2016**, *252*, 40–45, doi:10.1016/j.pscychresns.2016.04.003.

387. Hitti, F.L.; Siegelbaum, S.A. The hippocampal CA2 region is essential for social memory. *Nature* **2014**, *508*, 88–92, doi:10.1038/nature13028.

388. Dudek, S.M.; Alexander, G.M.; Farris, S. Rediscovering area CA2: Unique properties and functions. *Nat. Rev. Neurosci.* 2016, *17*, 89–102.

389. Rubenstein, J.L.R.; Merzenich, M.M. Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain Behav.* **2003**, *2*, 255–267, doi:10.1034/j.1601-183X.2003.00037.x.

390. Bozzi, Y.; Provenzano, G.; Casarosa, S. Neurobiological bases of autism–epilepsy comorbidity: a focus on excitation/inhibition imbalance. *Eur. J. Neurosci.* **2018**, *47*, 534–548, doi:10.1111/ejn.13595.

391. Yasui, S.; Tsuzaki, K.; Ninomiya, H.; Floricel, F.; Asano, Y.; Maki, H.; Takamura, A.; Nanba, E.; Higaki, K.; Ohno, K. The TSC1 gene product hamartin interacts with NADE. *Mol. Cell. Neurosci.* **2007**, *35*, 100–108, doi:10.1016/j.mcn.2007.02.007.

392. Samanta, D. An Updated Review of Tuberous Sclerosis Complex-Associated Autism Spectrum Disorder. *Pediatr. Neurol.* 2020, *109*, 4–11.

393. Lipton, J.O.; Sahin, M. The Neurology of mTOR. *Neuron* **2014**, *84*, 275–291, doi:10.1016/j.neuron.2014.09.034.

394. Switon, K.; Kotulska, K.; Janusz-Kaminska, A.; Zmorzynska, J.; Jaworski, J. Molecular neurobiology of mTOR. *Neuroscience* **2017**, *341*, 112–153, doi:10.1016/j.neuroscience.2016.11.017.

395. Eltokhi, A.; Gonzalez-Lozano, M.A.; Oettl, L.-L.; Rozov, A.; Pitzer, C.; Röth, R.; Berkel, S.; Hüser, M.; Harten, A.; Kelsch, W.; et al. Imbalanced post- and extrasynaptic SHANK2A functions during development affect social behavior in SHANK2-mediated neuropsychiatric disorders. *Mol. Psychiatry* **2021**, doi:10.1038/s41380-021-01140-y.

396. Nelson, R.J. The use of genetic "knockout" mice in behavioral endocrinology research. *Horm. Behav.* **1997**, *31*, 188–196, doi:10.1006/hbeh.1997.1381.

397. Grienberger, C.; Konnerth, A. Imaging Calcium in Neurons. *Neuron* **2012**, *73*, 862–885, doi:10.1016/j.neuron.2012.02.011.

398. Lin, M.Z.; Schnitzer, M.J. Genetically encoded indicators of neuronal activity. *Nat. Neurosci.* **2016**, *19*, 1142–1153, doi:10.1038/nn.4359.

399. Broussard, G.J.; Liang, R.; Tian, L. Monitoring activity in neural circuits with genetically encoded indicators. *Front. Mol. Neurosci.* **2014**, *7*, doi:10.3389/fnmol.2014.00097.

400. Antinucci, P.; Hindges, R. A crystal-clear zebrafish for in vivo imaging. *Sci. Rep.* **2016**, *6*, 1–10, doi:10.1038/srep29490.

401. Dooley, C.M.; Schwarz, H.; Mueller, K.P.; Mongera, A.; Konantz, M.; Neuhauss, S.C.F.; Nüsslein-Volhard, C.; Geisler, R. Slc45a2 and V-ATPase are regulators of melanosomal pH homeostasis in zebrafish, providing a mechanism for human pigment evolution and disease. *Pigment Cell Melanoma Res.* **2013**, *26*, 205–217,

doi:10.1111/pcmr.12053.

402. Romano, S.A.; Pérez-Schuster, V.; Jouary, A.; Boulanger-Weill, J.; Candeo, A.; Pietri, T.; Sumbre, G. An integrated calcium imaging processing toolbox for the analysis of neuronal population dynamics. *PLoS Comput. Biol.* **2017**, *13*, doi:10.1371/journal.pcbi.1005526.

403. Giovannucci, A.; Friedrich, J.; Gunn, P.; Kalfon, J.; Brown, B.L.; Koay, S.A.; Taxidis, J.; Najafi, F.; Gauthier, J.L.; Zhou, P.; et al. CaImAn an open source tool for scalable calcium imaging data analysis. *Elife* **2019**, *8*, 1–45, doi:10.7554/eLife.38173.

404. Pnevmatikakis, E.A.; Soudry, D.; Gao, Y.; Machado, T.A.; Merel, J.; Pfau, D.; Reardon, T.; Mu, Y.; Lacefield, C.; Yang, W.; et al. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **2016**, *89*, 285, doi:10.1016/j.neuron.2015.11.037.

405. Kirkby, L.A.; Sack, G.S.; Firl, A.; Feller, M.B. A role for correlated spontaneous activity in the assembly of neural circuits. *Neuron* 2013, *80*, 1129–1144.

406. Marachlian, E.; Avitan, L.; Goodhill, G.J.; Sumbre, G. Principles of functional circuit connectivity: Insights from spontaneous activity in the zebrafish optic tectum. *Front. Neural Circuits* **2018**, *12*, 1–8, doi:10.3389/fncir.2018.00046.

407. Avitan, L.; Pujic, Z.; Mölter, J.; Van De Poll, M.; Sun, B.; Teng, H.; Amor, R.; Scott, E.K.; Goodhill, G.J. Spontaneous Activity in the Zebrafish Tectum Reorganizes over Development and Is Influenced by Visual Experience. *Curr. Biol.* **2017**, *27*, 2407-2419.e4, doi:10.1016/j.cub.2017.06.056.

408. Lal, D.; Trucks, H.; Møller, R.S.; Hjalgrim, H.; Koeleman, B.P.C.; De Kovel, C.G.F.; Visscher, F.; Weber, Y.G.; Lerche, H.; Becker, F.; et al. Rare exonic deletions of the RBFOX1 gene increase risk of idiopathic generalized epilepsy. *Epilepsia* **2013**, *54*, 265–271, doi:10.1111/epi.12084.

409. Barbera, G.; Liang, B.; Zhang, L.; Gerfen, C.R.; Culurciello, E.; Chen, R.; Li, Y.; Lin, D.T. Spatially Compact Neural Clusters in the Dorsal Striatum Encode Locomotion Relevant Information. *Neuron* **2016**, *92*, 202–213, doi:10.1016/j.neuron.2016.08.037.

410. Cai, D.J.; Aharoni, D.; Shuman, T.; Shobe, J.; Biane, J.; Song, W.; Wei, B.; Veshkini, M.; La-Vu, M.; Lou, J.; et al. A shared neural ensemble links distinct contextual memories encoded close in time. *Nature* **2016**, *534*, 115–118, doi:10.1038/nature17955.

411. Mu, Y.; Bennett, D. V.; Rubinov, M.; Narayan, S.; Yang, C.T.; Tanimoto, M.; Mensh, B.D.; Looger, L.L.; Ahrens, M.B. Glia Accumulate Evidence that Actions Are Futile and Suppress Unsuccessful Behavior. *Cell* **2019**, *178*, 27-43.e19, doi:10.1016/j.cell.2019.05.050.

412. Lee, P.H.; Feng, Y.C.A.; Smoller, J.W. Pleiotropy and Cross-Disorder Genetics Among Psychiatric Disorders. *Biol. Psychiatry* 2021, *89*, 20–31.

413. Smoller, J.W.; Andreassen, O.A.; Edenberg, H.J.; Faraone, S. V.; Glatt, S.J.; Kendler, K.S. Psychiatric genetics and the structure of psychopathology. *Mol. Psychiatry* **2019**, *24*, 409–420, doi:10.1038/s41380-017-0010-4.

414. Gandal, M.J.; Haney, J.R.; Parikshak, N.N.; Leppa, V.; Ramaswami, G.; Hartl, C.; Schork, A.J.; Appadurai, V.; Buil, A.; Werge, T.M.; et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science (80-. ).* **2018**, *359*, 693–697, doi:10.1126/science.aad6469.

# ANNEX

# PAIRING HUMAN GENETIC FINDINGS WITH MODEL ORGANISMS - THE TRANSLATIONAL GENETICS OF ADHD AND ITS MENTAL COMORBIDITIES

Judit Cabana-Domínguez[a,b,c,d]*, Ester Antón-Galindo[a,b,c,d]*, Noelia Fernández-Castillo[a,b,c,d]*, Euginia Singgih[e]*, Aet O'Leary[fg], William HG Norton [h], Annette Schenck [e], David Slattery[f], Bru Cormand[a,b,c,d,@]

[a] Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Barcelona, Catalonia, Spain

[b] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Spain

[c] Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Catalonia, Spain

[d] Institut de Recerca Sant Joan de Déu (IRSJD), Esplugues de Llobregat, Catalonia, Spain

[e] Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, the Netherland

[f] Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital, Goethe University, Frankfurt, Germany

[g] Division of Neuropsychopharmacology, Department of Psychology, University of Tartu, Tartu, Estonia

[h] Dept. Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, LE2 3DL, UK

* Equally contributed

@ Corresponding author:

Bru Cormand: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 3ª planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4021013. Fax: (+34) 93 4034420; Email: bcormand@ub.edu

**To be submitted to Neuroscience and Biobehavioral Reviews**

## INDEX

## 1. Genetic animal models of ADHD and comorbidities

Attention deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder that affects approximately 5% of children and adolescents and 2.5% of adults worldwide. ADHD is markedly impairing, as it can significantly increase the risk for substance abuse and for other psychiatric disorders (about 89% of ADHD individuals have a comorbid psychiatric disorder [1] and contribute to educational and occupational failure, accidents, and criminality. ADHD results from the interaction of genetic and environmental risk factors that alter the structure and function of brain networks involved in behaviour and cognition. Twin studies have estimated a heritability around 70-80% [2], and genome-wide association studies (GWAS) estimated a SNP heritability that ranges from 0.10 to 0.28 [3–5], supporting the contribution of common variants to the etiology of ADHD. The largest GWAS of ADHD to date (20K ADHD patients and 35K controls) identified 12 independent risk loci, adding important new information about the underlying biology of ADHD [4]. Furthermore, several studies have demonstrated a significant genetic overlap between ADHD and other psychiatric disorders, such as autism spectrum disorder (ASD), major depression (MD) and schizophrenia [3,6], many of which co-occur frequently with ADHD. Although a lot of progress has been made during the last five years in defining the genetic landscape of ADHD, there is still a long way to go to fully understand the molecular underpinnings of the disorder. The use of animal models can help us to understand the role of specific genes related, not only to ADHD, but also to its comorbid traits. Here, we review the genes that have been related to ADHD and other comorbid psychiatric traits in rodents, zebrafish and *Drosophila*, as well as the different tests used to study these phenotypes and different pharmacological approaches applied (Table 1). Importantly, we highlight the strengths and limitations of the use of each animal model.

**Table 1.** Summary of tests performed to test ADHD-related phenotypes and its comorbid disorders in rodents, zebrafish and drosophila.

| Disorder | Traits | Tests used | | |
|---|---|---|---|---|
| | | Rodents | Zebrafish | Drosophila |
| ADHD-related symptoms | Hyperactivity | Open-field test, water maze | Locomotive assays | Activity monitoring, capillary feeder (CAFE) assay, open-field assay |
| | Impulsivity | 5-choice serial reaction time task | Locomotion (swimming) monitoring, 5-choice serial reaction time task | Courtship disinhibition assay |
| | Inattention | 5-choice serial reaction time task, visual detection task, virtual object recognition task | 5-choice serial reaction time task, object recognition task, social attention paradigm | Tethered flight paradigms, Buridan's paradigms, optomotor maze |
| Autism spectrum disorder | Impaired social behaviour and communication, stereotypic behaviour, cognitive rigidity | Three-chambered social approach, partition test, nesting behaviour, ultrasonic vocalizations, open-field test, Morris water maze, T maze | Shoaling assays, Y maze, interaction with conspecifics, visually-mediated social preference test | Habituation learning assay, grooming, social behavior assay, courtship song assay, Y-maze |
| Aggressive behaviour | Aggression, social dominance | Resident intruder test, Dyadic social interaction test, social dominance test | Dyadic fight test, interaction with mirror image assay | Dyadic fight test |
| Anxiety | Anxiety-related behaviours, thigmotaxis | Open field, elevated plus maze, elevated zero maze, light dark box | Active avoidance conditioning | Open-field assay |
| Major depression | Anhedonia, despair | Sucrose preference test, Porsolt forced swim test, Tail-suspension test | | Effect of learned helplessness upon chronic unpredictable mild stress in various readouts |
| Schizophrenia | Impaired sensorimotor gating | Prepulse inhibition test | Prepulse inhibition test | Larval prepulse inhibition test |
| Substance use disorders | Reward | Drug-induced locomotor activity or conditionated place preference | Place preference paradigm | Appetitive taste memory test, associative learning assay |

## 2. Rodents

Rodents have been extensively used in psychiatric research because of their sophisticated behavioural repertoire and their (relative) genetic similarity to humans. Moreover, while human brains are clearly larger and more developed, core anatomical features are shared between rodents and humans, including the structures and networks that govern particular behaviours. For example, in both humans and rodents, the fear and reward circuits are well-conserved [7].

Historically, rats were the preferred species in behavioural research due to their ability to quickly learn and perform complex cognitive tasks without much experimenter manipulation. However, from the 1980s until recently, the advance of genetic tools to manipulate the mouse genome led to an explosion in their use in preclinical settings and the conversion of rat tasks to those more suited to mice. The recent development of, for example knockout rats, CRISPR-CasP9, TALEN and RNAi technologies, has seen an increase in the use of genetically-manipulated rats [8]. While complex behaviours, such as attention and impulsivity cannot be assessed until after weaning, this period of the rodents development correlates well with adolescence in humans. Therefore, rodents can be used to look across most of the lifespan at behaviours related to ADHD and possible influences of genetic and drug treatments on these. Moreover, rodents can easily be used to determine the effects of a variety of environmental insults or enrichments concomitant with drug and genetic studies. In keeping with the general direction in psychiatric research, i.e., the RDOC approach [9], the complexity of ADHD cannot be fully replicated in preclinical models but specific traits or endophenotypes can be. Thus, investigation of such behaviours will ultimately provide greater understanding of the neurobiological bases of these traits and, by extension, the disorder as a whole.

## 2.1. Testing ADHD-related behaviours in rodents

A wide-range of behavioural tests and tasks can be used in mice and rats to assess phenotypes that resemble symptoms observed in ADHD patients. The main tests that are currently employed in preclinical research are listed in Tables 1-2; although this is not a comprehensive list since there are a number of tasks, such as the probabilistic reversal learning task [10] and the affective bias test [11], which have not yet been widely used in ADHD-related research. Below, we provide only a short overview of the three main phenotypic domains (hyperactivity, impulsivity and inattention) that clinically characterize ADHD and can be assessed in rodents. Both mice and rats are suitable for determining the underlying neurobiology and drug/genetic components that can bidirectionally affect these domains.

*Hyperactivity* is the easiest domain to quantify and can be monitored via automated tracking systems and software or even simply via counting line crossings. Moreover, different forms of activity can be recorded, such as home-cage locomotion, which can reveal changes across different phases of the circadian cycle, or novelty-induced locomotor activity in a novel environment (i.e. an open field chamber).

**Table 2.** Description of common tests used in rodents for assessing ADHD-like, and comorbid endophenotypes

| Behavioural test | Trait | Description | Comments |
|---|---|---|---|
| **5-CSRTT** | Attention / Impulsivity | Rodents are trained in operant chambers to respond to a brief visual stimulus presented in one of five spatial locations. The animal must respond to the correct location to receive food reward. Premature responding is a measure of impulsive action / response impulsivity. The test can be modified to the 5-choice CPT, which includes behavioural inhibition trials | Initially developed as a rodent analogue of CPT but does not require response inhibition to a presented stimulus. Training is time-consuming, relatively low throughput. Requires food restriction and may be biased by changes in motivation and attention. Responsive to methylphenidate and amphetamine. |
| **CPT** | Impulsivity | Rodents are trained in operant chambers to respond to a target stimulus and withhold a response to non-target stimuli. False alarm rate in this test provides an index of motor impulsivity. | Relatively labor-intensive, relatively low throughput. Requires food restriction and may be biased by changes in motivation and attention. Responsive to methylphenidate and amphetamine |
| **Stop-signal task** | Impulsivity | Rats are trained in operant chambers to nose-poke to initiate a trial and perform a "go" response for food reward. In 20% of the trials, an auditory stop-signal is presented at a predetermined time before the "go" response can be completed. Inhibiting completion of the response in "stop" trials is rewarded. | Relatively labor-intensive, low throughput. Requires food restriction and may be biased by changes in motivation. Responsive to amphetamine which abolishes striatal lesion induced deficits. |
| **Go / No-Go** | Impulsivity | Rodents are trained in operant chambers to respond differentially to two distinct cues. During a "go" cue (e.g., a light stimulus), a nose-poke into an active hole is rewarded. During a "no-go" cue (e.g., a tone), inhibition of the response into the active hole is rewarded. | This works better in rats than mice as it is difficult to train mice to initiate trials. Relatively labor-intensive, low throughput. Requires food restriction, especially in mice. |
| **Delay discounting** | Impulsivity | Rodents are trained to choose between a small reward available immediately or a large reward delivered after a variable delay. Impulsive animals are more negatively affected by the delay to large reward delivery. Measures impulsive choice / decision impulsivity. | Labor-intensive, low throughput. Requires food restriction and may be biased by changes in motivation. Responsive to psychostimulants which reduce the rate of delay discounting |
| **Open field** | Hyperactivity, anxiety | Rodents placed in an open field arena (circular or square) can investigate the peripheral zone (typically darker and with tactile stimulus from the walls) or the central zone. | Ease of use, high throughput, responsive to anxiolytics, inter-laboratory reliability, but strain-dependent and dependent on the light conditions could be considered more a test for basal locomotor activity than anxiety-related behaviour |

| | | | |
|---|---|---|---|
| **Differential reinforcement of low rates of responding task** | Impulsivity | Rats are trained in operant chambers with an active and inactive lever available. Active lever press results in food reward. Animals are gradually trained to inhibit active lever presses for up to 20s to receive a reward. Measures impulsive responding and impulsive choice. | Relatively labour-intensive. Requires food restriction and may be biased by changes in motivation. Responsive to psychostimulants. |
| **Variable delay-to-signal** | Impulsivity | Animals are trained to withhold responding to a light during a pre-determined delay period but are rewarded when responding to the same stimulus during the response period. Training to criterion concludes with a single test session in which half of the trials have a variable prolonged delay period and responding during the delay period is not punished. | Training is faster than for the better-known operant impulsivity paradigms. Measures of impulsive action and impulsive choice obtained within one test session. Good correlation with 5-CSRTT and DD measures. Responsive to methamphetamine. |
| **Home-cage activity** | Spontaneous locomotor activity | Animals are assessed either via camera tracking or via implanted microchips in their home-cage over long periods of time. This allows researchers to determine whether the rodents show altered circadian activity | Can be difficult to set up, although newer software with machine learning algorithms are making this easier to track multiple animals in one cage at a time. Implanted microchips have the advantage of providing extra information such as blood pressure, heart rate and body temperature |
| **Saccharine / sucrose preference** | Anhedonia | same as above but swimming and climbing behaviours are separated and are increased with serotonergic or catecholaminergic antidepressants, respectively. | Ease of use, high throughput, responsive to ADs and chronic stress models, inter-laboratory reliability, can be used repeatedly |
| **Forced swim test** | Depression-like behaviour, AD screening | Rodents, placed in an inescapable container of water swim more following antidepressant administration | Ease of use, high throughput, responsive to AD, inter-laboratory reliability, but responsive to acute drug administration, reliant on motor function and not always responsive to SSRIs |
| **Modified forced swim test** | Depression-like behaviour, AD screening | Same as above but swimming and climbing behaviours are separated and are increased with serotonergic or catecholaminergic antidepressants, respectively. | Ease of use, high throughput, responsive to AD (including SSRIs), inter-laboratory reliability, but responsive to acute drug administration and reliant on motor function |
| **Tail suspension test** | Depression-like behaviour, AD screening | Rodents, chiefly mice, when hung from the tail will adopt an immobile posture. Antidepressant treatment increases the time animals spend in active behaviours | Ease of use, high throughput, inter-laboratory reliability, responsive to antidepressants, but may differ from FST results, responsive to acute drug administration and reliant on motor function |
| **Female urine sniffing test** | Anhedonia, sexual motivation | Male rodents will prefer to investigate the sniff a cotton bud soaked in the urine from an oestrous female over water. Antidepressant treatment increases the time animals investigate the urine whereas chronic stress decreases it. Thought to be a measure of sexual motivation | The test is responsive to both stress and antidepressant treatment and can be performed with a relatively high throughput. However, while gaining in popularity, it has not been widely used and requires further studies to increase its validity |
| **Progressive ratio responding** | Reward sensitivity, motivation | Rodents can be trained to work for food or drug rewards and their motivation to work for the reward (i.e. lever pressing) assessed using a progressive ratio responding paradigm whereby the animal has to work harder and harder for each subsequent reward | Widely used to assess reward motivation, inter-laboratory reliability, but not high throughput and requires more validation |

| | | | |
|---|---|---|---|
| **Drug-induced locomotor activity** | Drugs effects | Open field test, locomotor activity boxes and other apparatuses can be used to study the acute and chronic effects of various drugs on general locomotor activity, as well as assessment for sensitisation after a withdrawal period. | Easy to use, high throughput. |
| **Conditioned place preference** | Reward | In repeated sessions, administration of a drug (or any other stimulus of interest) is paired with one specific chamber in one session, and vehicle is paired with another chamber in a separate session. In a probe test, animals can freely choose between the chambers. Increased time spent in the chamber paired with the drug indicates conditioned place preference and is a measure of the rewarding properties of the drug. | Easy to use, relatively low throughput. Addictive substances readily induce place preference, this test is frequently used to screen drugs for abuse potential. |
| **Elevated zero / plus maze** | Anxiety | A conflict-based test where are cross-shaped or circular maze with two open (brightly-lit) and two closed (darker and with walls for tactile information) assess a rodent's exploratory drive and anxiety of bright and open spaces. Less anxious animals will spend more time in the open arms | Ease of use, high throughput, responsive to anxiolytics, inter-laboratory reliability, but strain-dependent and locomotor based |
| **Light-dark box** | Anxiety | A conflict-based test were rodents can investigate a brightly lit compartment or a dark compartment. Another test examining a rodent's exploratory drive and anxiety of bright and open spaces. Less anxious animals will transition more between the zones and/or spend more time in the light box | Ease of use, high throughput, responsive to anxiolytics, inter-laboratory reliability, but strain-dependent and locomotor based |
| **Stress-induced hyperthermia** | Anxiety | Taking the body temperature of a rodent via a rectal thermometer represents a stressor and when assess again (15-30min) later there is a slight rise (typically $0.5 – 1.0^oC$) | Ease of use, high throughput, responsive to anxiolytics, a physiological alternative to other tests |
| **Vogel test** | Anxiety | Water deprived rodents will drink water and every set number of licks receive a mild shock | Responsive to anxiolytics, but not widely used and requires shocks to be effective, which requires stronger ethical approval and may be biased by changes in pain sensitivity |
| **Defensive burying** | Anxiety | Rats receive a shock when an object is placed into their cage and will thereafter attempt to bury it | Responsive to anxiolytics, but not widely used and requires shocks to be effective, which requires stronger ethical approval and may be biased by changes in pain sensitivity |
| **Four-plate test** | Anxiety | Each time a rodent moves from one plate to another, shaped in a 2x2 grid, they receive a mild footshock | Responsive to anxiolytics, but not widely used and requires shocks to be effective, which requires stronger ethical approval and may be biased by changes in pain sensitivity |

| | | | |
|---|---|---|---|
| **Social preference / memory** | Social-behaviour | Typically performed by letting a rodent investigate a novel arena for 3 min in which there is an empty target/interaction zone followed by a further 3 min during which a conspecific is confined within the target zone. The first phase of the test provides information about locomotor activity and general anxiety and the time ratio the experimental animal spends investigating the target zone (time with conspecific / time with empty target) provides a sociability index. If required, another phase can be added to assess social memory whereby the same conspecific and a novel conspecific are introduced in different locations and the time interacting with each assessed; if social memory is intact the experimental animal should favour interacting with the novel conspecific | Easy to use, relatively high throughput, widely used and responsive to a wide range of environmental and drug manipulations |
| **Social interaction** | Social-behaviour | Two conspecifics, of the same age/weight are placed into a novel arena and the time spent investigating is assessed. | Easy to use, relatively high throughput, widely used and responsive to a wide range of environmental and drug manipulations |
| **Three-chamber test** | Social-behaviour | Typically, there are three sessions in an apparatus that takes the form of three connected chambers. In the first, habituation phase, general activity is assessed and then for the second phase a conspecific (in a wire-mesh container) is added to one of the outer chambers. In the final phase, the same conspecific is added to one of the outer chambers and a novel conspecific to the other. The second phase measures general sociability/preference and the final stage assesses social memory as the experimental animal should favour interacting with the novel conspecific | Easy to use, relatively high throughput, requires some more specific equipment than other tests on the list, responsive to a wide range of environmental and drug manipulations |
| **Tube /Social dominance test** | Social-behaviour | This is a test of social dominance, partly through aggression, in mice. Two mice are introduced to a clear plastic tube, one at each end and when they meet in the middle the dominant / more aggressive conspecific will force the other to reverse out of the tube | Easy to use, relatively high throughput, not as widely used as some of the other tests, but a useful test to determine social dominance / hierarchy |
| **Resident intruder** | Social-behaviour | A test to assess aggressive behaviour in rats and mice. Typically, the experimental male is housed together with a female, to increase territorial aggressive behaviour and a smaller conspecific is introduced to the home-cage (the female is removed for the time of testing). This can be performed repeatedly across a number of different sessions to observe an escalation in aggressive behaviour | Easy to use, relatively high throughput, if the resident displays too much aggressive behaviour the test has to be terminated. Can also be performed in a neutral arena |

| | | | |
|---|---|---|---|
| **Partition test** | Social-behaviour | This test is conducted in a standard cage which is divided in two by a wire mesh or perforated plastic partition wall. The experimental mouse is able to see, hear, and smell, but not physically interact with the stimulus mouse placed behind the partition. Time spent close to the partition is an index of social interest. | Easy to use, relatively high throughput. Can be modified to assess social preference/memory by introducing novel/familiar social stimuli |
| **Nesting behaviour** | Self-care, Wellbeing | Rodents readily build nests if suitable material is available. For nesting behaviour, a pre-weighed cotton square is placed into the home cage of the experimental mouse and the quality of the nest is assessed 24h later. Poor quality nests indicate reduced self-care and compromised wellbeing. | Easy to use, high throughput. |
| **Ultrasonic vocalizations** | Social behaviour | Adult rodents, including mice, consistently emit a variety of ultrasonic vocalisations (USVs) during various types of social interactions. These USVs, recorded via ultrasonic microphones, are a measure of social interest and motivation. | The recording of USVs is well-established for rats but is less common in mice. Mice emit USVs during male-female and female-female encounters but not male-male interactions. |
| **T/Y-maze spontaneous alternation test** | Spatial working memory / Spatial novelty preference | To assess spatial working memory, mice are placed in one arm of the T/Y-maze, typically for 5 min, and allowed to move freely through the arms of the maze. Entries into each arm are recorded, and a successful alternation is defined as sequential entries into all three arms. The proportion of spontaneous alternations out of all arm entries is calculated. | Easy to use, high throughput. Can be conducted as a spontaneous or a forced/baited task. |
| **Morris water maze** | Spatial learning and memory | A mouse is placed into a large circular water tank, filled with opaque 25 °C water (mixed with milk or other substance), and required to locate a submerged platform. The escape latency and length of path to platform is measured in repeated sessions. This is a test for spatial learning/re-learning and memory. | Relatively easy to use, although relatively low throughput. Can be used to assess initial spatial learning, as well as re-learning by moving the platform to another location. |
| **Prepulse inhibition test** | Impaired sensorimotor gating | Sensorimotor gating is assessed by measuring the reduction of the acoustic startle response to a startling stimulus (typically 120dB white noise) induced by a weak sub-threshold pre-stimulus. | Easy to use, relatively high throughput. |

Abbreviations: AD, antidepressant; USV, ultrasonic vocalisation

*Impulsivity and inattention* can be looked at using operant tasks, which can be performed in traditional Skinner boxes or more sophisticated touchscreen chambers. Regardless of equipment, the tasks evaluated in the lever-pressing/nose-poke apparatus and in the touchscreen are similar. In some cases, the touchscreen version can include more complex visual stimuli and allow for more fine-tuned and subtle parameters to be manipulated. However, whether this mechanism alters the behaviour of the rodent has not been extensively looked at. Also, various paradigms are available in rodents to determine the individual's ability to perform set-shifting task [12], whereby the reward can be identified either via the substrate (i.e. sawdust vs bedding) or the scent (i.e. ginger vs mint). Similarly, both intra- and inter-dimensional set-shifting can be evaluated. 5-choice serial reaction time task (5-CSRTT) can be performed in rats and mice, with some differences in training between the species, and allow researchers to determine a variety of behavioural components such as impulsivity, (in)attention, perseverance and motivation [13]. The continuous performance test [14,15] has recently been back-translated into a versatile test for assessing sustained attention, behavioural inhibition and motivation in rodents. Another test that has recently been developed for use in rodents is delay discounting, which is mainly used to evaluate impulsive choice, as well as to determine motivation and attention [16,17]. As with the clinical version of the test, the rodent has to choose between a small immediate reward and a larger delayed reward.

In Table 2, we have also described tests that can be performed in rodents to assess endophenotypes related to ADHD comorbidities. Since these tests are not directly associated to ADHD and have been extensively reviewed previously, we refer the reader to the table and recent reviews [7,18–20].

## 2.2. Strains and transgenic rodent lines used as ADHD models

In rodents, some spontaneous mutations lead to a hyperactive phenotype. It is the case of Coloboma (Cm) mice [21] and the Spontaneous Hyperactive Rat (SHR) [22]. Cm is a mouse strain developed by neutron irradiation that caused a mutation on chromosome 2 which disrupts about 20 genes (recently reviewed [23]). This mutation includes *Snap25*, pointed as the putative causal gene of the phenotype since polymorphisms in this gene have been associated with ADHD in humans [24,25] and the rescue of its expression in Cm mice reduces hyperactivity. Cm mice display the three main features of ADHD: inattention, impulsivity and hyperactivity, and the last one seems to be attenuated with amphetamine but not with methylphenidate [21]. On the other hand, the SHR strain was developed by inbreeding of rats of the Wistar-Kyoto (WKY) strain.

These rats also show the core symptoms of ADHD (hyperactivity, impulsivity and inattention) compared to WKY rats [26,27]. However, these results are controversial [28], especially because WKY is a particularly inactive strain [29]. Besides, the most classical neurodevelopmental model of ADHD created by lesion in brain systems was obtained by neonatal 6-hydroxydopamine (6-OHDA) injection. These mice exhibit the major ADHD-like symptoms (hyperactivity, attention deficit and impulsivity) together with comorbid behaviours that usually appear in ADHD patients, such as anxiety-like and antisocial behaviours and decreased cognitive functioning [30].

Several knockout and transgenic mice have been proposed as ADHD models, mostly based on targeting genes involved in dopamine transmission, a key neurotransmitter in ADHD. One of the best described ADHD mouse models is the dopamine transporter (*Dat* or *Slc6a3*) knockout (KO) mouse [31]. DAT plays a critical role in regulating extracellular dopamine concentration and strong evidence supports that abnormal DAT function may be involved in ADHD. Indeed, both increased and decreased *DAT* expression have been reported in human ADHD patients [32–34] and genetic studies have reported associations between gene variants and ADHD [35]. The *DAT* knockdown mouse (*Dat*-KD) [36] or the *DAT* cocaine-insensitive mouse (*Dat*-CI) [37,38] are other *Dat* mutant mice presenting ADHD-related features. Finally, transgenic mouse models not targeting dopaminergic genes also exist, such as the tachykinin-1 (*Nk1*) receptor KO mouse [39,40] and the Trβpv knock-in (KI) mouse [41,42], that present altered monoaminergic transmission and ADHD core behavioural features.

## 2.3. Mouse models to study ADHD and its comorbidities: Inspecting the Jackson database

Since not much information is available on mouse models to study ADHD together with other comorbidities, we explored available information on different mouse lines to identify novel models that could be used for ADHD, either with or without other psychiatric comorbid phenotype. To systematically browse all the existing mouse lines and strains that present ADHD-related behaviours we used the Jackson Laboratory mouse strains database (https://www.jax.org/mouse-search). We performed a search of mouse strains that present any of the three main ADHD-related phenotypes: "hyperactivity", "impulsivity" or "inattention". We found 172 transgenic strains with specific genetic modifications in single genes (Supplementary Table S1). We then reported their behavioural alterations and found that, interestingly, 111 of these strains present only ADHD-related behaviours and 62 others present also behaviours related to other psychiatric disorders (Tables 3-4 and Supplementary Table S1). To further

understand the functions of these genes we performed a KEGG pathway and Gene Ontology (GO) Biological Process over-representation analysis (www.webgestalt.org).

### 2.3.1. ADHD without comorbidities

In the Jackson database we found 111 transgenic mouse strains (with a total of 103 genes mutated) that present exclusively one or more of the ADHD-related phenotypes (hyperactivity, impulsivity or inattention), and not any additional behaviour related to other psychiatric disorders (Table 3 and Supplementary Table S1). From these 111 strains, 104 present only hyperactivity, one presents only impulsivity (*Comt*) and three strains present only inattention (*Psen1*, *Snap25* and *Tardbp*) (see Supplementary Table S1). The identification of these strains is likely biased to hyperactivity probably because this behaviour can easily be evaluated in a simple open-field test, frequently used as a routine test, whether attention and impulsivity have to be tested with a specific test such as 5-CSRTT (Table 1-2) [43], or can also be evaluated with other visual detection tasks requiring attentional engagement. In addition, several complex tests might not have been performed for some of these strains, so these results could be incomplete or biased, and these animals could present other psychiatric phenotypes not tested yet. We performed a gene-based association study using summary statistics from the last ADHD GWAS meta-analysis [4] and found that 11 out of the 103 genes altered in these strains with exclusively ADHD-related behaviours were associated to ADHD (Table 3). The gene-based analysis was performed with MAGMA (v1.06) [44] using the 1000 Genomes Project Phase 3 (Build 37/European data only) as a reference panel.

**Table 3.** Genes related to ADHD phenotypes identified in transgenic mouse lines (Jackson database)

| | Traits | Test used | Genes | KEGG pathways* | GO Biological Process* | Genes associated with ADHD in the gene-based analysis[1] |
|---|---|---|---|---|---|---|
| **ADHD-related symptoms (with and without comorbidities)** | Hyperactivity | Open-field test, water maze | *Abca2; Abcg1; Actl6b; Adcy3; Adcyap1; Adipor2; Ankfn1; Anks1b; Ap3b2; Ap3d1; Apaf1; App; Arrdc3; Arsa; Atf2; Atp1a3; Atrn; Bdnf; Cacna2d3; Cacna2d4; Cacng2; Cadm1; Calm1; Camk2a; Cdh23; Cdk17; Cdk5r1; Cdkl5; Celf4; Chd3; Chd7; Chrd; Chrm1; Chrm4; Cic; Ckap5; Clic5; Cntnap2; Comt; Crebbp; Dgat1; Dgkb; Disc1; Dnajb5; Drd1; Drd2; Drd3; Dtnbp1; Dusp18; Eef1b2; Elmod3; En2; Eps15l1; Espn; Esr1; Fmr1; Fos; Foxi1; Fxr2; Gabra1; Gabra3; Gabrb3; Git1; Glra1; Gnai2; Gnao1; Gpr135; Gpr88; Gpx6; Gria1; Grid2; Grin2b; Hmox1; Htr2c; Htt; Igsf9b; Il6; Ints3; Kcna4; Kcne1; Ldlr; Lepr; Lmx1a; Lrrk2; Magi2; Maob; Mapk3; Mapt; Mcoln3; Myo6; Myo7a; Ncor1; Nlgn2; Nlgn3; Nox3; Npas3; Npc1; Nr4a2; Nr4a3; Nup153; Oprd1; Otc; Otog; Per1; Pitx3; Pkd2l2; Pnpla6; Pou4f3; Ppargc1a; Ppfia3; Ppm1f; Psap; Psen1; Ptchd1; Ptprk; Rab5b; Rgs4; Rnf214; Scn1a; Shank2; Shank3; Sirt1; Slc12a6; Slc1a2; Slc26a10; Slc5a7; Slc6a3; Slc6a8; Slc9a6; Snai2; Snap25; Snca; Sobp; Syngap1; Syt4; Tardbp; Tbc1d8; Tbx10; Tecpr2; tip; Tmie; Uba6; Ush1c; Ush1g; Vim; Vldlr; Wdr41; Whrn; Zbtb20; Zeb1; Zpld1* | Dopaminergic synapse Amphetamine addiction cAMP signaling pathway Circadian entrainment Adrenergic signaling in cardiomyocytes Neuroactive ligand-receptor interaction Glutamatergic synapse Aldosterone synthesis and secretion Alzheimer disease Lysosome | Regulation of trans-synaptic signaling Regulation of membrane potential Cognition Locomotory behavior Synapse organization Neuron death Neurotransmitter transport Ear development Divalent inorganic cation transport Response to radiation | *ADCY3, ARRDC3, ARSA, ATP1A3, BDNF, CELF4, GLRA1, GNAI2, GPX6, GRIA1, GRID2, MAPT, MYO7A, NPAS3, OPRD1, PPFIA3, RGS4, TMIE* |
| | Impulsivity | 5-choice serial reaction time task | *Cadm1; Comt; Per1; Shank3* | | | |
| | Inattention | 5-choice serial reaction time task, visual detection task, virtual object recognition task | *Comt; Psen1; Ptchd1; Snap25; Tardbp* | | | |

| | | | Gene list | Pathways | GO terms | Genes¹ |
|---|---|---|---|---|---|---|
| **Exclusively ADHD-related symptoms (without comorbidities)** | Hyperactivity | Open-field test, water maze | *Abca2; Abcg1; Actl6b; Adcy3; Adcyap1; Adipor2; Ankfn1; Ap3b2; Ap3d1; Apaf1; APP695; Arrdc3; Arsa; Atf2; Atrn; Bdnf; Cacna2d4; Cacng2; Calm1; Cdh23; Cdk17; Cdk5r1; Celf4; Chd3; Chd7; Chrd; Chrm4; Ckap5; Clic5; Dgat1; Dgkb; Dnajb5; Dusp18; Eef1b2; Elmod3; Eps15l1; Espn; Fos; Foxi1; Gabra1; Git1; Glra1; Gnai2; Gpr135; Gpx6; Grid2; Htr2c; Snca; Ints3; Kcna4; Kcne1; Ldlr; Lepr; Lmx1a; Lrrk2; Maob; Mapk3; Mcoln3; Myo6; Ncor1; Nox3; Nr4a3; Nup153; Otc; Otog; Per1; Pitx3; Pkd2l2; Pnpla6; Pou4f3; Ppfia3; Ppm1f; Psap; Ptchd1; Ptprk; Rab5b; Rtl10; Rxylt1; Sirt1; Slc12a6; Slc1a2; Slc26a10; Slc6a8; Slc9a6; Snai2; Snca; Sobp; Tbc1d8; Tbx10; Tecpr2; tip; Tmie; Ush1c; Ush1g; Vldlr; Wdr41; Whrn; Zeb1; Zpld1* | Adrenergic signaling in cardiomyocytes<br>Cocaine addiction<br>Circadian entrainment<br>Amphetamine addiction<br>Alzheimer disease<br>Dopaminergic synapse<br>Cushing syndrome<br>Cholinergic synapse<br>Lysosome<br>Insulin secretion | Ear development<br>Locomotory behavior<br>Regulation of trans-synaptic signaling<br>Neuron death<br>Response to ammonium ion<br>Divalent inorganic cation transport<br>Forebrain development<br>Response to antibiotic<br>Response to oxidative stress<br>Organic hydroxy compound metabolic process | *ADCY3, ARRDC3, ARSA, BDNF, CELF4, GLRA1, GNAI2, GPX6, GRID2, PPFIA3, TMIE* |
| | Impulsivity | 5-choice serial reaction time task | *Comt; Per1* | | | |
| | Inattention | 5-choice serial reaction time task, visual detection task, virtual object recognition task | *Comt; Psen1; Ptchd1; Snap25; Tardbp* | | | |

* Identified using WebGestalt software. Pathways and GO terms sorted by significance of enrichment, FDR<0.05. Weighted set cover was applied to reduce redundancy.

¹ Genes presenting a nominal association (p-value<0.05) in the gene-based analyisis using summary statistics from Demontis et al., 2019.

**Table 4.** Genes related to ADHD phenotypes and its comorbid disorders identified in transgenic mouse lines (Jackson database)

| Comorbidity | Traits | Test used | Genes | KEGG pathways* | GO Biological Process* | Genes previously related to patients | |
|---|---|---|---|---|---|---|---|
| | | | | | | ADHD | Comorbid disorder |
| **Autism spectrum disorders** | Impaired social behaviour and communication, stereotypic behaviour, cognitive rigidity | Three-chambered social approach, partition test, nesting behaviour, ultrasonic vocalizations, open-field test, Morris water maze, T maze | *Anks1b, Cdkl5, Cntnap2, Crebbp, Disc1, En2, Fmr1, Gabrb3, Gnao1, Gria1, Grin2b, Htt, Magi2, Mapt, Nlgn2, Nlgn3, Npas3, Rgs4, Scn1a, Shank2, Shank3, Syngap1,Uba6* | Nicotine addiction, Long-term potentiation, Glutamatergic synapse, Dopaminergic synapse | Regulation of synapse structure or activity, Localization within membrane, Cognition, Locomotor behaviour, Regulation of membrane potential, Neuron death | *Anks1b, Cdkl5, Cntnap2, Disc1, Fmr1, Grin2b, Nlgn2, Npas3, Rgs4, Syngap1, Uba6* | *Anks1b, Cdkl5, Cntnap2, Crebbp, Disc1, En2, Fmr1, Gabrb3, Gria1, Grin2b, Htt, Mapt, Nlgn2, Nlgn3, Scn1a, Shank2, Shank3, Syngap1* |
| **Aggressive behaviour** | Aggression, social dominance | Resident intruder test, Dyadic social interaction test, social dominance test | *Gria1, Cacna2d3, Cadm1, Camk2a, Crebbp, Disc1, En2, Esr1, Fmr1, Rgs4* | Long-term potentiation | - | *Camk2, Esr1, Fmr1, Rgs4* | *Disc1,Esr1, Fmr1,Rgs4* |
| **Anxiety** | Anxiety-related behaviours, thigmotaxis | Open field, elevated plus maze, elevated zero maze, light dark box | *Atp1a3, Cadm1, Camk2a, Cdkl5, Cic, Crebbp, Drd3, Fmr1, Gria1, Igsf9b, Il6, Magi2, Mapt, Myo7a, Nlgn2, Npc1, Oprd1, Rnf214, Shank2, Shank3, Slc5a7, Syngap1, Syt4, Uba6, Vim, Zbtb20* | - | Adult behavior, Cognition, Neuromuscular process, Synapse organization, Regulation of cell morphogenesis, Regulation of neurotransmitter levels | *Atp1a3,Cdkl5, Cic, Drd3, Fmr1, Il6, Shank3, Slc5a7, Syngap1, Uba6, Zbtb20* | *Atp1a3, Drd3, Fmr1, Il6, Nlgn2* |
| **Major depression** | Anhedonia, despair | Sucrose preference test, Porsolt forced swim test, Tail-suspension test | *Camk2a, Il6, Magi2, Nr4a2, Ppargc1a, Shank3, Syt4,* | - | Glutamate receptor signaling pathway, Response to xenobiotic stimulus, Positive regulation of neuron differentiation | *Il6, Nr4a2, Ppargc1a,* | *Il6, Nr4a2, Ppargc1a,* |
| **Schizophrenia** | Impaired sensorimotor gating | Prepulse inhibition test | *Anks1b, Fxr2, Gabra3, Hmox1, Npas3, Shank3, Syngap1* | - | Localization within membrane | *Anks1b, Npas3, Syngap1* | *Anks1b, Fxr2, Hmox1, Npas3, Shank3, Syngap1* |
| **Substance use disorders** | Reward | Drug-induced locomotor activity or conditionated place preference | *Atp1a3, Cadm1, Cic, Chrm1, Drd2, Drd1, Drd3, Dtnbp1, Gpr88, Gria1, Nr4a2, Shank3, Slc6a3* | Dopaminergic synapse, cAMP signaling pathway | Locomotory behavior, Cognition, Synapse organization | *Drd1, Drd2, Drd3, Nr4a2, Slc6a3* | *Drd1, Drd2, Drd3, Dtnbp1, Slc6a3* |

* Identified using WebGestalt software. Pathways and GO terms sort by significance of enrichment, FDR<0.05. Weighted set cover was applied to reduce redundancy.

Interestingly, we found that this gene set of 103 altered mouse genes presenting exclusively ADHD-related behaviours is enriched for genes involved in regulation of trans-synaptic signalling and forebrain development, as well as for dopaminergic synapse genes, among other categories (Table 3). There is ample evidence at genetic, pharmacological, neuroimaging and neuropsychological levels that dysregulation of synaptic transmission and dopaminergic pathways contributes to the pathophysiology of ADHD [45–48]. Thus, it is expected that animal models with altered dopaminergic genes present ADHD-related behaviours, as it has been reported in *DAT* KO mice.

### 2.3.2. ADHD and comorbidities

A striking feature of ADHD clinical manifestation is the frequent co-occurrence with other neuropsychiatric conditions [49]. Up to 89% of individuals with ADHD also receive a diagnosis of one or more additional psychiatric disorders (Sobanski, 2006). We subsequently explore rodent models that present ADHD features together with other behavioural abnormalities related to typical ADHD comorbid major psychiatric conditions.

*ADHD and autism-spectrum disorder (ASD)*

Both ADHD and ASD are two early-onset neurodevelopmental disorders with a high comorbidity. Indeed, 20-50% of children with ADHD also meet criteria for ASD and genetic studies have demonstrated shared heritability and genetic overlap between these disorders [6,50,51]. ASD is defined by a range of social and communication deficit severities coupled with repetitive and unusual sensory-motor behaviours [52]. In mice, behavioural tests are used to evaluate traits resembling ASD features: impairment of social interaction and communication, repetitive behaviours and behavioural inflexibility (Table 1-2)[53–55]. In the Jackson database, we have found 27 transgenic mouse lines that present both hyperactivity and ASD-related symptoms such as impaired sociability or repetitive motor behaviours (Table 4 and Supplementary Table S1). From these mouse lines, 23 present impaired sociability, that was determined using the three-chambered social approach or the partition test by measuring decreased social interaction with littermates or reduced preference for investigating a social stimulus (Table 1-2) [55]. Repetitive behaviours are ASD-related traits that can be evaluated in an open field test identifying repetitive motor movements and increased self-grooming, or in a marble-burying test (Table 1-2) [55,56]. From these mouse lines, 16 also show increased stereotypic behaviours (Table 4 and Supplementary Table S1).

Added to the impaired sociability and repetitive behaviours, some mice present other traits that can be considered as ASD-related: decreased vocalization, which reveals communication impairments, cognitive rigidity, or altered nest building behaviour (Table 1-2). Indeed, the last one, a form of homecage-activity often linked to social behaviour [55], appears in nine transgenic mice (*Cdkl5 cKO, Cdkl5-/y, Cntnap2-/-, Crebbp-/-, Gabrb3-/-, Magi2 Tg, Shank2-/-, Shank3-/-, Uba6 NKO*). Another ASD trait present in some mice is cognitive rigidity and perseveration, that can be evaluated in a classic reversal task using the Morris water maze or the spontaneous alternation T maze test, as well as in a marble burying task (Table 1-2) [55,56]. *Cntnap2−/−* mice showed significantly higher number of no alternations in a standard T maze test and perseveration in the reversal task on the Morris water maze [57]. Also, *Fmr1^{I304N}* mice showed perseverative behaviour in the marble burying test [58]. Finally, a decrease in the number of ultrasonic vocalizations (USVs), normally emitted by mice in social situations, has been related to communication deficits, relevant to ASD. Ultrasonic vocalizations are emitted by pups separated from the dam and nest, during juvenile interactions, by resident females in a resident–intruder task and by males responding to female urinary pheromones (Table 1-2) [55,56]. *Cntnap2-/-* pups and *Shank3* Tg (conditional KI that results in overexpressed *Shank3* in dendritic spines in the cortex, hippocampus and striatum) pups emitted significantly lower number of ultrasonic calls than WT littermates [57,59]. Also, when allowed to interact with a novel WT female mouse, *Shank2−/−* male mice emitted ultrasonic vocalizations less frequently than did WT animals, and took longer to make the first call [60]. In a pup retrieval assay, *Shank2−/−* female mice retrieved the pups less efficiently than did WT mice [60].

This group of 27 genes that are altered in mouse lines showing hyperactivity and ASD-related symptoms is enriched in genes that participate in cognition, regulation of synapse structure and activity, and locomotor behaviour, important functions related to ASD and ADHD. The majority of the genes altered in these transgenic mouse lines are listed in the SFARI database (https://gene.sfari.org/) as genes implicated in the susceptibility to autism, and some of them were also related to ADHD in patients (Table 4 and Supplementary Table S2). These genes are highly expressed in brain, where they play a role in synaptic transmission, cell adhesion or neurogenesis. From these mouse lines, seven could be used specifically for investigating ADHD and ASD, without other reported behavioural alteration (transgenic lines for *Cntnap2, Gabrb3, Gnao1, Grin2b, Htt, Nlgn3* and *Scn1a*).

*ADHD and aggressive behaviour*

Aggressive behaviour is highly comorbid with ADHD and can be assessed as a trait or as part of diagnostic categories such as conduct disorder, opponent-defiant disorder or callous unemotional. About 47% of children with ADHD have opponent-defiant disorder and around 30% to 50% of them have comorbid conduct disorder [61]. Conversely, ADHD prevalence is also high in young and adult offenders, estimated around 30% [62,63]. Moreover, a recent GWAS meta-analysis identified three genome-wide significant loci for ADHD and disruptive behaviour disorders, suggesting a shared genetic architecture between these disorders [64].

In animal models, the most widely used test to assess aggression is the resident-intruder, since aggression often occurs in mice to establish and defend a territory, measured by the number of attacks or bites and the latency to attack (Table 1-2) [65]. Social dominance can be assessed using the tube test. In general, aggressive behaviour is only tested in males, and when assessed in females usually maternal aggression is tested [66]. In the Jackson database we found 10 transgenic mouse lines presenting hyperactivity and altered aggressive behaviour, which showed either increased aggression (*Cacna2d3-/-*, *Cadm1 Tg*, *Camk2a+/-*, *Disc1 Tg*), decreased aggression (*Crebbp Tg*, *Esr1-/-*, *En2-/-*, *Gria1-/-*) or decreased social dominance (*Fmr1-/-* and *Rgs4-/-)* (Table 4 and Supplementary Table S1).

Four of these genes have been related to aggressive behaviour in humans: *CAMK2A*, *ESR1*, *FMR1* and *RGS4* (Supplementary Table S2). Remarkably, polymorphisms in the *ESR1* gene, coding for the oestrogen receptor 1, have been associated with anger, neuroticism, indirect aggression and antisocial behaviour [67], and with ADHD [68]. Moreover, two knockout lines can be used to investigate ADHD and aggression comorbidity specifically, *Cacna2d3-/-* with increased aggression and *Esr1-/-* with decreased aggression.

*ADHD and anxiety*

Around 25% of children with ADHD have comorbid anxiety disorders [69,70]. The tests used in anxiety include elevated plus maze, elevated zero maze and light-dark box, in which a long time spent in closed or dark areas is associated with increased levels of anxiety (Table 1-2) [71]. In the open-field test, a high degree of thigmotaxis is associated with increased anxiety. Hyperactivity and differences in anxiety are present in 27 mouse transgenic lines in the Jackson database, the vast majority assessing anxiety in the elevated plus maze (Table 4 and Supplementary Table S1). From these lines 11 show increased anxiety-related behaviours (*Cic-/-*, *Il6-/-*, *Magi2 Tg*, *Myo7a sp*,

*Nlgn2 Tg, Oprd1-/-, Shank2-/-, Shank3-/-, Slc5a7 Tg, Uba6-/-, Vim-/-*), and 16 lines show decreased anxiety-related behaviours (*Atp1a3 Tg, Cadm1 Tg , Camk2a+/-, Cdkl5-/y , Crebbp Tg, Drd3-/- , Fmr1 Tg, Gria1-/-, Igsf9b-/- , Mapt Tg , Npc1 Tg, Rnf214-/-, Syngap1+/-, Syngap Tg, Syt4-/-, Zbtb20 Tg).* Remarkably, sex differences were identified for one of them: *Magi2* transgenic mice show increased anxiety in males but not in females [72]. It should be considered that anxiety in *Camk2+/-* and *Uba6-/-* lines was only assessed using the open-field test, in which the animals showed increased thigmotaxis, but no other specific tests for anxiety were used [73–75]. In addition to hyperactivity, *Cadm1 Tg* and *Shank3-/-* mouse lines showed also impulsive behaviour [76,77]. These genes seem to be involved in synapse organization, regulation of neurotransmitter levels, cognition and adult behaviour.

From these genes, 11 have been previously related to ADHD in patients, and *ATP1A3, DRD3, FMR1, IL6, NLGN2* with anxiety (Table 3 and Supplementary Table S2). Interleukin-6, encoded by *IL6* gene, is involved in inflammatory response, and higher IL6 serum levels are found in patients with ADHD [78–80] and anxiety [81,82] . Four mouse lines could be used to test specifically ADHD with comorbid anxiety (increased levels), *Myo7a sp, Oprd1-/-, Slc5a7 Tg*, and *Vim-/-*. From those, the transgenic line for *Slc5a7* is of special interest, since this gene has previously been related to ADHD [83].

### ADHD and major depression (MD)

ADHD and MD are two psychiatric conditions that co-occur frequently, being ADHD 7.5 times more prevalent in chronic MD than in the general population [84]. Importantly, genetic studies have demonstrated the existence of shared genetic risk factors between them using different bioinformatic approaches [6,85].

In the Jackson database we found seven transgenic mice strains that present both hyperactivity and alterations in depressive-like behaviour (Table 4 and Supplementary Table S1). One of the core symptoms of depression is anhedonia, the reduced ability to experience pleasure, which can be tested in rodents with the sucrose preference test (Table 1-2). We found anhedonia in two transgenic strains, *Ppargc1a-/-* mice gene [86,87],and a knock-in for *Magi2* gene expressed under the control of *Camk2a* in the excitatory neurons of the forebrain [72]. On the other hand, *Il6-/-* mice showed enhanced hedonic behaviour in the sucrose preference test [72,88]. Another methodological approach to test despair, depression-like behaviour in mice, and to evaluate the efficacy of antidepressant drugs is the Porsolt forced swim test (Table 1-2) [89]. In this test, the *Nr4a2+/-* mice showed a depression-like profile compared to WT animals [90]. Conversely,

reduced despair and depression-like behaviour was found in *Camk2a-/-* [75], *Syt4-/-* [91] and *Il6-/-* mice [92]. In the case of *Il6-/-* mice, these results are supported by the ones obtained in the sucrose preference test [88]. Finally, another indicator of despair in mice is the immobility during the tail-suspension test. *Shank3 Tg* mice showed a reduction in the duration of immobility compared with WT, suggesting a reduction in despair in the transgenic animals [59].

These genes seem to participate in the glutamate receptor signalling pathway, response to xenobiotic stimulus and neuron differentiation. Interestingly, some of those genes have previously been associated both with ADHD and MD, such as *PPARGC1A*, *NR4A2* and *IL-6* (Supplementary Table S2). From these transgenic lines, *Ppargc1a-/-* is the only one not showing other comorbid behavioural alterations.

*ADHD and schizophrenia*

Recent genetic studies have demonstrated a significant genetic correlation between ADHD and schizophrenia [3,6]. Moreover, both disorders share symptoms such as impaired attention and deficits in inhibition and working memory [93]. In mice, impaired prepulse inhibition (PPI) is widely accepted as the most significant endophenotype of schizophrenia and it is considered indicative of disrupted sensorimotor gating, which clinically correlates in patients with symptoms such as thought disorder and distractibility (Table 1-2) [94–96]. In the Jackson database we have found eight transgenic mice lines that present both hyperactivity and decreased PPI and involve seven different genes (*Anks1b*, *Fxr2*, *Gabra3*, *Hmox1*, *Npas3*, *Shank3*, *Syngap1*) (Table 2 and Supplementary Table S1). Two other transgenic mice strains presented increased PPI (involving *Fmr1* and *Mapt* genes), but this altered behaviour is not related to schizophrenia. Individuals with schizophrenia suffer also from various cognitive deficits, including impaired working memory, that can be tested in mouse models using different paradigms such as novel object recognition, contextual and cued fear conditioning, or mazes (Table 1-2) [94]. In the Jackson database, three mice lines (*Shank3 Tg*, *Syngap1+/-*, *Fxr2-/-*) present also working memory impairments added to decreased PPI, and five mouse lines (*Anks1b-/-*, *Npas-/-*, *Shank3 Tg*, *Shank3-/-*, *Syngap1+/-*) present impaired social behaviour, another endophenotype of schizophrenia [94], added to a decreased PPI.

All these genes, except for *GABRA3*, have already been related to schizophrenia in patients (Table 2 and Supplementary Table S2). Moreover, microdeletions in *ANKS1B*, de novo mutations in *SYNGAP1* and common variants in *NPAS3* were found in ADHD patients [97–99]. Three of these

transgenic lines, for the genes *Fxr2*, *Gabra3* and *Hmox1*, show only alterations related to ADHD and schizophrenia with the tests performed so far.

### *ADHD and Substance Use Disorders (SUD)*

Apart from the core ADHD symptoms, this psychiatric disorder is associated with an increased risk of harmful outcomes like substance abuse, and about 40% of ADHD patients present lifetime SUD [100]. This high co-occurrence could be explained both by environmental [101,102] and genetic risk factors [85,103]. We have found 12 transgenic mice strains in the Jackson database that present hyperactivity and alterations on substance use phenotypes (Table 4 and Supplementary Table S1). The effects of drugs of abuse in mice can be studied using both unconditioned and conditioned behaviours [104]. Regarding the first ones, the most studied is the drug-induced locomotor activity (Table 1-2), apparently produced by increased dopamine release. We inspected the transgenic mice models that showed alterations in locomotor activity in response to drugs. Compared with WT animals, *Chrm1-/-* [105] and autoDrd2-/- mice (lack of D2 autoreceptors specifically) [106] display supersensitivity to locomotor effects of cocaine. Similarly, other mouse strains exhibit locomotor supersensitivity to drugs: *Gria1-/-* mice after morphine injection [107,108], and *Nr4a2+/-*, *Atp1a3(Myk/+)*, *Chrm1-/-*, *Gpr88-/-* and *Shank3 Tg* mice after amphetamine administration [59,90,105,109,110]. Conversely, a low-dose of amphetamine in the conditional *Cic-/-* mice (deletion of Cic in the developing forebrain) and in *Cadm1 Tg* mice (GFAP-DNSynCAM1, dominant-negative form of SynCAM1 specifically targeted to astrocytes) exerted a paradoxical calming effect [111], previously described in patients and some ADHD mice models [111]. Interestingly, *Dat*-CI mice showed increased locomotion induced by amphetamine and morphine, but not by cocaine [111], and *Drd1-/-* mice showed a cocaine dose-dependent decrease in locomotion [112].

On the other hand, the rewarding effects of drugs can also be studied using conditioned behaviours like conditioned place preference (CPP) o drug self-administration (Table 1-2). We have found two transgenic mice that showed hyperactivity and alterations in CPP. *Drd3-/-* mice showed increased morphine-induced CPP at lower doses compared to WT, but this effect was attenuated at the highest dose [113,114]. *Dat-CI* mice were unable to develop CPP induced by cocaine, but amphetamine produced similar levels of CPP than in WT [37], suggesting that the lack of response was produced only by the inability of cocaine to block DAT. Finally, *Dys−/−* mice showed hyperactivity combined with alterations in the operant learning paradigm (self-administration) with reward pellet, that may be due to increased impulsive and compulsive

behaviour during early sessions [115,116]. However, these deficits can be improved with increased training and experience.

Some of the genes highlighted here are involved in the dopaminergic (*Drd1*, *Drd2*, *Drd3* and *Slc6a2*) or glutamatergic (*Gria1* and *Dtnbp1*) neurotransmission systems, key elements of the reward system and widely studied both in ADHD and drug addiction (Table 2 and Supplementary Table S2). Six of these mouse lines could be used to test specifically ADHD with comorbid SUD, in particular the transgenic lines for *Chrm1*, *Drd1*, *Drd2*, *Dtnbp1*, *Gpr88* and *Slc6a3* (*Dat*), being the genes *Drd1*, *Drd2* and *Slc6a3* previously related to ADHD and SUD in patients (Supplementary Table S2).

### 2.3.3. Strains involving several genes

In the Jackson database we have found four transgenic and six spontaneous or radiation-induced mouse strains with mutations in more than one gene that display hyperactivity among other behavioural and neurological phenotypes (Supplementary Table S3). Apart from the Cm mice, described above, the most interesting ones are the mice strains that bear a deletion that mimics a copy number variation (CNV) on human 16p11.2. This CNV encompasses 26 genes that are highly conserved on mouse chromosome 7F3, it has been reported in ADHD patients and it is among the most common genetic variations found in ASD [117]. There are two mouse strains with an heterozygous deletion of this region: 16p11+/- mice exhibit normal social behaviour but show hyperactivity [118], and 16p11.2df mice show both hyperactivity and stereotypic behaviours [119]. On the other hand, we found mice with a duplication of about 3Mb (about 19 genes) in chromosome 11 (Dp(11)17) that spans the genomic interval commonly deleted in Smith-Magenis syndrome patients, who present a behavioural phenotype that closely resembles ADHD [120]. Those mice show hyperactivity together with abnormal social interaction and increased anxiety [121].

## 2.4. Models used to test pharmacological treatments

As stated above, rodents are ideal to investigate genetic and environmental factors underlying ADHD-related phenotypes and the effect of ADHD medications on them. Indeed, in the last few decades, numerous studies have investigated how candidate gene deficiency affects the behavioural traits related to the core and comorbid symptoms of ADHD. The most extensively studied gene in this context in mouse studies is the one coding for the dopamine transporter (DAT), given the fact that psychostimulants acting predominantly via this transporter represent

the gold standard of pharmacological treatment. The consensus from these studies is that DAT deficiency causes novelty-induced hyperactivity, increased impulsivity, inattention, and cognitive impairments. These deficits can be reversed by both amphetamine and methylphenidate (for a review, [122]. These findings have been corroborated in the recently generated DAT-KO rats [123], thus supporting the involvement of dopamine transporter in ADHD-related phenotypes.

Candidate-gene association studies and GWAS have identified a number of novel ADHD risk genes and, as mentioned above, several transgenic mouse models have been generated to investigate their potential role in ADHD pathogenesis. One of them is *ADGRL3 (*previously named *LPHN3)*, knockout mice for this gene are hyperactive, impulsive, and display increased social behaviour and decreased aggression. While the effect of methylphenidate and amphetamine have not yet been published, *Adgrl3*-deficiency was shown to dysregulate cortical DAT expression [124]. These findings were replicated in a rat model, as the *Adgrl3-KO* rats were shown to have a blunted response to amphetamine [125]. Another candidate gene, *CNTNAP2*, has been associated with ADHD and ASD and its deletion in mice results in hyperactivity and social impairments [57,126]. Interestingly, risperidone attenuates the hyperactivity but does not rescue the social deficits. For more details on these studies we refer the interested reader to our recent review paper which describes cross-species (drosophila, zebrafish, mouse, and human cell lines) findings on these genes [127].

The first twelve genome-wide significant risk loci for ADHD were recently described [4]. Some of these genes have previously been linked to behavioural alterations. It is the case of *FOXP2,* that plays a well-established role in speech and language and has previously been associated with schizophrenia [128–130]. *Foxp2* knockout mice present major deficits in reversal learning together with a downregulation of D1R expression [131], and abnormal social behaviour [132]. Also, a risk variant in *Tmem161b* (rs10514299) has been shown to predict striatal activation during reward processing in alcohol dependence [133], which is further evidence that this target is worth pursuing in a mouse model. Finally, while no behavioural findings have been reported for *Dusp6-/-* (dual-specificity phosphatase 6) mice, this gene has been related to obesity, which is an often-reported comorbidity of ADHD [134]. Even though the majority of these risk genes have not been previously linked to ADHD-related behavioural changes, they represent attractive targets for further exploration.

## 2.5. Concluding remarks - Rodents

To our knowledge, no models had been proposed to date for the study of ADHD with comorbid conditions. Our systematic searches in the Jackson database have provided new insights on different mouse lines that show primarily hyperactivity with or without alterations in other behaviours related to psychiatric comorbid conditions. As mentioned before, the selection of these mouse lines is biased to hyperactivity since this behaviour is easily tested in routine tests, such as the open-field test. Moreover, although less commonly assessed, it is also possible to study whether this observed hyperactivity is due to novelty and/or lack of habituation (i.e. open-field) compared with basal (home-cage) activity. Although numerous tests for measuring inattention and impulsivity traits are available, they have only been tested and described in few mouse lines in the database due, in part, to the requirement of specialised equipment and active, as opposed to passive, measurement. Thus, it would be interesting to perform specific tests assessing inattention and impulsivity (Table 1-2) in these mouse lines to investigate the presence of these ADHD-related traits apart from hyperactivity. The same should be mentioned regarding other behavioural alterations, as specific tests like PPI, CPP or resident-intruder test, might have not been performed in many mouse lines. Additionally, it is important to mention that hyperactivity can be a potential confound in the interpretation of findings from other tests, since the majority rely on locomotor activity. In tasks such as the 5-CSRTT or CPT hyperactivity does not affect the results, but in other behavioural tests, especially in tasks that have a longer duration and varying stimulus presentation times, hyperactivity might be a confounding factor.

Most of the mouse lines that were identified in our review feature alterations in only one gene and the majority involve genes that have been previously related to ADHD and comorbidities in patients (Table 3-4 and Supplementary Table S2), which make them good models to use in future studies to investigate shared genetic mechanisms by ADHD and comorbid disorders. Moreover, as described above [4], the recent discovery of 12 genome-wide hits for ADHD represent strong targets to assess in genetically-modified mice. Nevertheless, we should consider that a single-gene approach does not capture the full genetic complexity of psychiatric disorders, although it could be a starting point to unravel the underlying neurobiological mechanisms.

## 3.  Zebrafish

The zebrafish has been used as a model for developmental biology for decades because of their rapid development and transparency at embryonic stages. Zebrafish develop outside of the mother making it easy to collect and manipulate embryos. Tools to manipulate genes, ablate cells and both visualise and manipulate neural activity using light have also been established [135–137]. In parallel, robust behavioural tests have been set up, enabling zebrafish to be used in translational studies of human diseases including psychiatric disorders [138]. Although the formation, position, and function of neurotransmitter signalling pathways sometimes differ between zebrafish and other vertebrates, comparative studies are beginning to precisely map these differences, allowing the transfer of information between species [139]. The ease of generating large numbers of zebrafish make them ideal for high-throughput analyses and imaging studies. As a model for translational studies, the zebrafish is particularly useful for optogenetic dissection of the behaviour, time-lapse analysis of neural development and screens for novel therapeutic treatments. The ability to apply chemical compounds to fish by immersion rather than injection into the stomach or brain makes zebrafish an excellent animal for screens to identify psychoactive drugs, an approach that has already been used for aggression, sleep and feeding [140–142]. Despite the impossibility of fully modelling a complex disorder such as ADHD, zebrafish have already been used to study different aspects of this disease. In the next sections we will describe recent research into the neurobiology of ADHD in this species.

### 3.1. Testing ADHD-related behaviours in zebrafish

*Hyperactivity.* Hyperactivity is the easiest of the three core symptoms of ADHD to measure in zebrafish. Zebrafish larvae move around consistently from about 5 days onwards, and at around one month, at the time of sexual maturation, zebrafish undergo a metamorphosis in the shape of the body and fins, altering the drag forces that act upon the body (Green and Hale, 2011). This gives rise to the adult pattern of locomotion in which contraction of trunk muscles is used to propel the body forwards[143,144]. Activity can be quantified by placing a single fish in a tank and using videotracking to extract parameters such as distance swum, speed of swimming, number of movement bouts and acceleration within bouts (Table 1) [141]. Previous reports of zebrafish ADHD-like models have described both an increase in the distance swum [145–147] and heightened acceleration during swim bouts, termed motor impulsivity [145,148]. However, changes to locomotion are a fairly non-specific read-out of fish behaviour – more evidence is required to relate this phenotype to ADHD. In the case of *adhesion G protein-coupled receptor L3.1*,

*period1b* and *micall2b,* such hyperactivity could be rescued by applying ADHD treatment drugs such as methylphenidate, atomoxetine and deprenyl, suggesting that hyperactivity can be used as an endophenotype for ADHD [145–147].

*Impulsivity.* Two types of impulsivity have been described in zebrafish, both of which may be useful to investigate the neurobiology of ADHD: motor and cognitive impulsivity. Motor impulsivity, as described above, represents sharp bouts of acceleration followed by periods of inactivity in contrast to the smooth locomotion curve usually displayed by larval zebrafish [145,148]. However, whether these periods of acceleration really represent impulsivity has been questioned [149] and further research is required to understand this phenotype. Cognitive impulsivity can be measured using the 5-CSRTT, a sophisticated test that has been adapted from a similar rodent paradigm (Table 1-2). A single adult fish is placed into an arena that contains five light emitting diodes (LEDs). The fish first selects an illuminated LED by nose poking to collect a food reward. Once this has been learnt, a variable inter-trial interval is added in which animals have to wait several sections before nose poking the LED. Selection of an incorrect (non-illuminated) LED can be scored as reduced attention (an error of omission), whereas inability to wait for the duration of the inter-trial interval is scored as impulsivity [149,150]. Pre-treatment of zebrafish with atomoxetine or amphetamine decreases impulsivity in this task, whereas methylphenidate has the opposite effect, making fish more impulsive [150]. This suggests that noradrenaline signalling may influence performance of this task in fish. A simpler version of this test, using only two choice arenas, has been used to characterise *period1b* mutant zebrafish, which displayed inattention in this task compared to wild-type animals [146]. This suggests that the 5-CSRTT may be a useful test for other ADHD models, even though it can only be used at adult stages.

*Inattention.* Changes in attention are a core symptom of ADHD that are likely to involve a number of physiological and psychological processes. To date, there have been few studies that have measured attention in zebrafish, and it is not clear whether zebrafish can maintain an attention set in a similar manner to other vertebrates [151]. Major categories of attention include: orienting, expectancy, stimulus differentiation, sustained attention, and parallel processing [152]. Orienting has been measured in a social attention paradigm in which male zebrafish were permitted to eavesdrop upon different stimuli: two male zebrafish fighting each other; two non-interacting males separated by a barrier; or an empty tank [153] (Table 1). The focal fish's orientation and proximity to the stimulus was used as a read-out of attention. Zebrafish spent more time watching a fight than watching the two fish separated by a barrier. They paid

particular attention to lateral displays at the beginning of fights suggesting that this represents an important social cue [153]. In a recent study, the ability of zebrafish to orient themselves to each other has been shown to be controlled by *LIM homeobox 8a*-positive cholinergic neurons in the ventral forebrain [154], an area of the teleost brain that may be homologous to the lateral septum in mammals. Sustained attention has been measured using a novel object recognition test [155] (Table 1). The amount of time spent interacting with an object presented on a video screen was recorded. Wild-type zebrafish could differentiate between a familiar and novel object up to 24 hours later. 5-CSRTT has also been used to quantify sustained attention (Table 1) [156]. Zebrafish have to select one of five same colour LEDs to get a food reward, ignoring the non-illuminated stimuli. Although they are capable of performing this task, zebrafish have a lower accuracy rate and response time on this test compared to rodents. This means that it is not clear whether attention measured in the 5-CSRTT can really be compared between zebrafish and other vertebrates. Zebrafish also appear to be capable of maintaining an attentional set, the ability to apply a set of rules to complex stimuli in order to differentiate relevant from irrelevant cues. This was measured by reversal learning coupled to an intra-dimensional set shift (for example, a change in the colour of a shape). Adult wild-type fish were first trained to pair one of two colour cues with a food reward. Once they had learnt this task the rule was swapped so that the second colour was now paired with food. A second set of experiments introduced an intra-dimensional set shift: the presentation of two new colours followed by a second rule change (once again switching the colour cue that led to reward). The amount of time that adult zebrafish took to learn these rules decreased as the experiment progressed suggesting that they had formed an attentional set and that the rule governing the correct response has been learnt [157]. There is also some evidence that zebrafish can undertake parallel processing in the form of feature binding - the ability to integrate the separate features of an object into the correct whole. The reaction of zebrafish to films of fish that varied in form and motion has been recorded [158]. Interestingly, focal fish spent more time in close proximity to congruent films (normal body shape and forward motion) compared to incongruent films (manipulated body shape; abnormal motion) indicative of feature binding. Feature binding has previously only been shown in mammals with a more developed neocortex. This study suggests that the teleost forebrain is capable of carrying out this task, meaning that the neural circuits required to perform feature binding may be simpler than originally thought [158].

## 3.2. Characterising ADHD-linked genes in zebrafish

Despite the difficulty of modelling all aspects of ADHD, zebrafish still represent an ideal model to investigate the expression and function of ADHD-linked genes in the brain. Knock-down or mutagenesis techniques can be used to investigate the function of candidate genes during neural development and the signalling pathways that they influence. Several ADHD candidate genes have been characterised in zebrafish, including *adhesion G protein-coupled receptor L3.1* (*adgrl3.1*) and *period1b*.

One of the first ADHD-linked genes to be studied in zebrafish was *adgrl3.1,* one of two zebrafish homologues of human *Adhesion G-protein Coupled Receptor 3*. *ADGRL3* risk variants increase the prediction of disease severity, long-term outcome and a patient's response to stimulant medication [159,160]. Many of the polymorphisms in *ADGRL3* associated with ADHD are located in introns rather than the coding region, suggesting that enhancers for other genes may be affected rather than *ADGRL3* itself [161]. Arguing against this, an enhancer called ECR47 is expressed in the ventral forebrain, midbrain and hindbrain of zebrafish in a manner that recapitulates some of the expression pattern of *adgrl3.1* [145,161]. Transient knock-down of *adgrl3.1* makes zebrafish larvae hyperactive during both the day and night [145,162]. These animals also alter their swimming trajectory, displaying sharp bouts of acceleration that have been called motor impulsivity [145]. Both of these phenotypes can be rescued by applying the prototypical ADHD treatment drugs methylphenidate and atomoxetine [145], a form of construct validity that suggests hyperactivity could be used as an endophenotype in this model. Reduction of *adgrl3.1* function also leads to a reduction and displacement of dopaminergic neurons in the posterior tuberculum, a part of the ventral diencephalon that is important for locomotion [163]. Individual posterior tuberculum neurons project both anteriorly and posteriorly with the majority of projections going to the spinal cord, suggesting homology with mammalian hypothalamic A11 DA neurons [163]. Although *adgrl3.1* morphants display similar levels of dopamine in the brain as wild-type siblings, they are insensitive to drugs that interact with D1 and D2-like dopamine receptors [164]. This suggests that morphants have a saturating increase in dopamine signalling that may increase locomotion by activating post-synaptic receptors. The link between *adgrl3.1* function and ADHD-like changes in behaviour has been confirmed in mouse and rat [124,125,165,166], providing strong evidence for the link between this gene and disrupted dopaminergic signalling. The ability to apply drugs to larval zebrafish by immersion means that zebrafish lacking *adgrl3.1* function represent an ideal model to screen for novel ADHD

treatments. This would require a stable mutant line to be created, and ideally, further behavioural phenotyping (including impulsivity and attention) to be carried out at adult stages.

The *period1b* (*per1b*) gene is part of the circadian clock that maintains diurnal rhythms. Circadian dysfunctions are thought to contribute to the aetiology of many psychiatric disorders, including ADHD. Adult zebrafish *per1b* mutants display hyperactivity, inattention in a two choice serial reaction time task (similar to the 5-CSRTT described above, see Table 1-2) and impulsivity [146], although this was measured in a mirror test that is frequently used to study aggression [167] (Table 1). They also have a disruption in their circadian changes in locomotion. The hyperactivity phenotype can be rescued with methylphenidate and deprenyl, adding weight to *per1b* mutants as a model for ADHD [146]. *Per1b* mutants display a reduction and disorganisation of posterior tuberculum dopamine neurons (similar to the *adgrl3.1* phenotype) as well as changes in the expression of genes related to dopamine signalling: *monoamine oxidase* (*mao*) and *dopamine beta hydroxylase* (*dbh)* expression is increased, whereas *orthopedia a, orthopedia b, mesencephalic astrocyte-derived neurotrophic factor* (*manf*), *wingless and integrated* (*wnt*) *1, wnt3a, wnt5a1* and *adgrl3.1* expression is decreased [146]. Excitingly, this suggests that both *per1b* and *adgrl3.1* act in a similar pathway to control ADHD-like behaviours via dopamine neurotransmission. Treatment of *per1b*-/- mutants with auriculasin, a prenylated isoflavone extracted from the root of *Flemingia philippinensis,* decreases hyperactivity and normalises the expression of dopamine-pathway genes [168]. This shows that auriculasin may represent a novel treatment option for some aspects of ADHD, and demonstrates the power of fish models to identify novel drug treatments. The *per1b* mutant is the most extensively characterised zebrafish ADHD model to date, mainly because it is a stable mutant line meaning that impulsivity and inattention can be measured in adult fish. However, the interpretation of the impulsivity phenotype may need investigating in more detail since altered interaction with a mirror could indicate decreased aggression in these mutants rather than inattention.

The Mical family is a conversed group of cytosolic multidomain proteins that are important for synaptogenesis, axon guidance and myofilament organisation. Polymorphisms in *MICALL2* were identified by a genome wide association study and eQTL sequencing of Han Chinese patients that display impaired executive inhibition, one of the core symptoms of ADHD [147]. *micall2b* (microtubule associated monooxygenase, calponin and LIM domain containing like protein 2b), one of the zebrafish homologues of *MICALL2,* is expressed in the nervous system, whereas *micall2a* is not [147]. Morpholino knock-down of *micall2b* triggers hyperactivity in larval zebrafish, a phenotype that can be rescued with atomoxetine [147]. However, neither attention nor

impulsivity has been measured in these animals, and the response of morphants to methylphenidate has not been measured [147]. This suggests that *micall2b* needs to be investigated in more detail, in particular to understand the role of this gene in nervous system development and synaptogenesis.

The RAB6A GEF Complex Partner 1 (Ric1) protein is important for collagen trafficking from the Golgi apparatus through the cell. Human patients with polymorphisms in *RIC1* display CATIFA syndrome that includes cleft lip, cataract, tooth abnormalities, intellectual disability, facial dysmorphism and ADHD [169]. *ric1-/-* mutant zebrafish exhibit reduced locomotion, a reduced forebrain and cerebellum, as well as a craniofacial phenotype and changes to the musculature [169]. Some of these phenotypes – such as the reduced forebrain and cerebellum size – may represent endophenotypes for ADHD. However, the reduction of locomotion, and lack of information regarding attention and impulsivity means that the link between this mutant line and ADHD is not very clear.

Several other ADHD-linked genes have been studied in zebrafish without considering their behavioural function. For example, a single nucleotide polymorphism in the last intron of *GFOD1 (Glucose-Fructose Oxidoreductase Domain Containing 1)* has been associated with inattention in a study of ADHD families [170]. In zebrafish, *gfod1* is widely expressed in the zebrafish nervous system, both during development and in adult animals [171]. There is prominent expression in GABAergic neurons, suggesting that this signalling pathway could contribute to the ADHD phenotype in human patients. In a similar approach, *SIRBP1* (Signal Regulatory Protein B1) has been identified by CNV analysis of patients with a high level of impulsive-disinhibition behaviour [172]. Expression analyses in zebrafish show that the fish *SIRPB1* homologue is expressed in the midbrain and muscle tissue. Further work would be required to understand if and how this contributes to impulsivity and ADHD [172].

### 3.3. Using zebrafish to investigate comorbid symptoms of ADHD

As mentioned above, human ADHD patients display a range of comorbidities with other psychiatric symptoms, including aggression, ASD and SUD. Zebrafish have been used to investigate behaviours linked to all of these disorders.

Aggression can be measured by either observing the behaviour of two fish in a dyadic fight, or recording a single animal's interaction with its own mirror image (Table 1) [173]. Zebrafish display characteristic agonistic postures that include extending their fins, thrashing the tail and

attempting to bite an opponent [167]. Zebrafish have been used to investigate novel genes that can influence expression of this behaviour [174] and to screen for drugs that decrease aggression [175].

Similar to ADHD, ASD is a complex psychiatric disorder that cannot be fully modelled in zebrafish. Although some ASD-linked symptoms can be studied, such as social interactions, other symptoms such as language impairment cannot be investigated in this model (Table 1) [176]. The zebrafish is a social species, and interactions with conspecifics can be investigated by characterising either shoaling, or the interaction with a single or several conspecifics placed in a tank compartment [177,178]. Several studies have characterised the expression of genes linked to ASD in the brain, and examined their behavioural function [179–181].

Zebrafish have also been used to measure reward behaviour, a component of drug addiction that could help understand SUD (Table 1). For example, a conditioned place preference (CPP) paradigm has been used to both identify changes in gene expression caused by exposure to amphetamine [182]. In a similar approach, a screen for mutant fish lines that display altered CPP behaviour has been used to identify *SLIT3* as an important gene mediating nicotine preference in both zebrafish and humans [183].

Together, these approaches demonstrate the power of zebrafish models to investigate a range of behavioural endophenotypes of psychiatric disorders. However, to date there are no studies that have measured comorbidities seen in human patients in a zebrafish ADHD model.

### 3.4. The effect of ADHD treatment drugs on neural development and behaviour

Current treatment options for ADHD include methylphenidate and atomoxetine, psychostimulants that interact with dopamine and noradrenaline neurotransmission. However, the effect of long-term stimulus medication on the developing brain has not been explored in detail. The zebrafish is an ideal model organism for such studies, because drugs can be applied by immersion. The effect of both acute and chronic methylphenidate treatment has been examined in wild-type zebrafish [184,185]. Acute immersion in 50 mg/L methylphenidate for the first five days of development increased the level of dopamine, noradrenaline and serotonin (5-HT) in the brain during larval stages [185]. However, this phenotype recovered by one month, with drug-treated animals showing similar levels of these neurotransmitters as untreated control siblings. Furthermore, methylphenidate exposed larvae spent more time at the bottom of a novel tank (an anxiety phenotype) and exhibited decreased choice accuracy in a spatial learning test at

adult stages, despite being drug-free for most of their lives [185]. This suggests that exposure to methylphenidate during gestation can have long term effects, even if the neurochemical phenotype recovered quickly. In a follow up study, Brenner and colleagues applied methylphenidate sub-chronically from 14, 21 or 28 days post fertilisation until 12 weeks before measuring anxiety, predator avoidance and social interaction. Methylphenidate tended to decrease the fish's response to environmental stimuli, and the drug had a stronger effect on behaviour when applied at later stages [184]. This is an interesting study that gives some insight into the possible effect of methylphenidate on children, and in particular on non-ADHD patients who might use this compound as a cognitive enhancer. It would be particularly insightful to repeat this research using a zebrafish ADHD model such as *adgrl3.1* or *per1b*, thereby adding more weight to the relevance of these findings to the disease.

## 3.5. Investigation of toxins and drugs that may trigger ADHD in humans

The symptoms of ADHD are caused by a combination of genetic susceptibility and environmental factors. Several studies have used zebrafish to examine how exposure to drugs or toxins during development can alter behaviour, with potential implications for this disease. Environmental exposure to lead can induce ADHD in human patients [186]. Zhang and colleagues treated zebrafish embryos with a low concentration of lead and observed a decrease in the number of axons forming tracts during development [187]. There were also parallel changes to the expression of key developmental genes including *eph receptor a4b*, *netrin 1b*, *netrin 2* and *no isthmus*. However, the axon tract phenotype recovered by 30 hours post fertilisation, and no changes to behaviour were investigated. This study demonstrates that lead exposure can have an impact upon neural development, but the link to ADHD is not that clear. In a similar approach, zebrafish have been immersed in either polychlorinated biphenyls (PCBs) or the perfluorinated compound perfluorooctane sulphonate [148,188]. Both types of pollutant have been linked to an increased incidence of ADHD in humans. Embryonic exposure to the PCB mixture Aroclor (A) 1254 enhanced thigmotaxis (the preference for the side of arena, used as a read-out of anxiety) and decreased the response to a visual startle stimulus that could be a measure of attention [188]. Treatment with perfluorooctane caused persistent hyperactivity and disorganised spontaneous locomotion including fewer bouts of swimming (suggestive of motor impulsivity). This phenotype could be rescued with dexamphetamine, indicating that PCB exposure could trigger ADHD in human patients [148]. These studies represent an interesting approach to determine how environmental effects can lead to the expression of this disease. However, in each case inclusion of a genetic model in the study (such as *adgrl3.1*, *per1b*) would provide more valuable

information about how genetic burden combines with environmental insult to contribute to psychiatric disease.

Zebrafish have also been used to investigate whether pain medication can lead to ADHD-like symptoms. Acetaminophen (also known as N-acetyl-p-aminophenol (APAP) or paracetamol) is a commonly used over-the-counter pain medication. There is evidence that prolonged acetaminophen use during pregnancy may increase the likelihood of a child displaying ADHD [189]. However, application of a low dose of acetaminophen during embryogenesis does not alter the locomotion of either wild-type of *adgrl3.1* morphants at 6 days [162]. Similar results were found in a study performed in wild-type mice [190]. This suggests that acetaminophen usage during gestation may not trigger ADHD, but further studies are needed to investigate this in more detail.

## 3.6. Concluding remarks - Zebrafish

Zebrafish have already been used as a model to investigate some aspects of ADHD, including the expression of candidate genes in the brain and the neurotransmitter signalling pathways they influence. Two interesting observations have arisen from these studies. Firstly, although the ADHD candidate genes *adgrl3.1* and *per1b* are widely expressed throughout the brain, loss of function appears to only affect a select group of dopamine neurons in the diencephalon [145,146]. It would be fascinating to understand why these neurons are more susceptible to loss of gene function compared to other groups. Secondly, both *adgrl3.1* and *per1b* appear to act in the same signalling pathway, and the posterior tuberculum neurons are very important for the control of locomotion. A more detailed characterisation of this brain area is needed to understand how it relates to groups of dopamine neurons in other vertebrate species.

Despite progress in establishing zebrafish as a model for this disease, there are some areas of research that could be improved. The current measurements of hyperactivity and (cognitive) impulsivity are convincing, but attention is still understudied in this species. The only zebrafish ADHD model that has been characterised for all three core symptoms of ADHD – hyperactivity, impulsivity and inattention – is *per1b* [146]. Other well-characterised candidate genes, such as *adgrl3.1* [145,164,191], need to be better characterised to investigate further aspects of this disease.

Some of the main strengths of fish as a model are the ability to measure behaviour at larval, juvenile and adult stages, and the ease with which drug screens can be carried out in this model. A comparison of how behaviour and neurobiology change over time could be particularly

informative for studies of ADHD, since some patients only express this disease during adolescence, whereas others first display symptoms at adult stages [192]. Several studies have used complex behavioural readouts to screen for drugs that control behaviour [140,142,175]. This approach could be leveraged to identify novel treatments for ADHD – although such an approach will likely need to use larval animals in order to increase throughput. One way to overcome this problem would be to first screen for drugs that can reduce hyperactivity at larval stages. A second round of screening could involve applying these drugs to adult fish and measuring impulsivity and perhaps inattention, followed by confirmation of a drugs effect in mouse. The zebrafish has already contributed to understanding of several human psychiatric disorders [138]. Coupled to the increasing number of tools available in this species, it seems likely that zebrafish are poised to increase our understanding of ADHD, including searching for novel drug treatments for this disease.

## 4.  *Drosophila Melanogster*

The fruit fly *Drosophila melanogaster* is a popular model in neurogenetics and has been used to establish the link between genes and behaviours for half a century [193,194]. *Drosophila* is cost-efficient and has relatively short generation time (~10 days). Approximately 75% of human genes have *Drosophila* equivalents [195]. The simple *Drosophila* genome is less redundant than the human one, and consequently mutations in a gene may cause more prominent phenotypes. The nervous system of *Drosophila*, with 15,000 neurons at larval stage and 250,000 in adulthood [196], is a relatively simple yet sufficiently complex model to study nervous system development, function, and behaviour. Importantly, while there is little similarity between human and *Drosophila* brain anatomy, the principle building blocks and many neuronal processes and mechanisms are conserved [197–199].

*Drosophila* is well suited to study neurodevelopmental and neuropsychiatric disorders, as it provides a multitude of approaches to investigate their underlying mechanisms and associated pathologies, from a molecular, subcellular, and circuit level to disease-relevant behaviour and cognitive processes. Such approaches include genetic or pharmacological induction of disease models, the former in a tightly spatiotemporally controlled manner if desired [200,201]. The generated models can then be phenotyped and/or molecularly characterized, then, phenotypes of interest can be further subjected to modification attempts, e.g. in genetic interaction and/or drug rescue experiments. The pool of publicly available stocks that can be readily utilized to

manipulate any gene of interest is enormous and steadily increasing [202–207]. Here, we will point out the main phenotypic assays that have been and can be applied in *Drosophila* models of ADHD (Table 1). Furthermore, we will highlight the insights that the characterization of ADHD risk genes in *Drosophila* has already provided. Lastly, we briefly discuss the potential and perspectives of using *Drosophila* as a disease model in this field.

## 4.1. *Drosophila* behavioural assays relevant for ADHD

ADHD is a phenotypically complex disorder, which in its complexity cannot be recapitulated in *Drosophila* or other animal models. Nonetheless, there are many behavioural traits sharing biological mechanisms with ADHD that lend themselves for animal modelling, and numerous approaches to study them in *Drosophila* exist [197,199,208]. This also applies to behavioural assays that are relevant to the core symptom domains of ADHD: (in)attention, (hyper)activity, and impulsivity.

*Hyperactivity.* Among the behavioural analyses, those quantifying locomotor activity and sleep appear highly relevant to characterize ADHD genes. Locomotor activity in the fly can model the hyperactivity component of the ADHD clinical diagnosis. Human ADHD genes were found to be enriched among *Drosophila* genes unbiasedly reported with ADHD face-valid behaviours including locomotor hyperactivity [209]. Locomotor activity is classically monitored in *Drosophila* using Trikinetics' well-established *Drosophila* Activity Monitoring (DAM) System, which registers the number of infrared beam crossings of individual flies at one or multiple positions of a test tube (Table 1). Recently, video tracking-based methods have increasingly gained interest due to their higher resolution and ability to assess different locomotor components. Video-based tracking also allows assessment of additional states and behaviours such as arousal, sleep pressure, and feeding [eg. *Drosophila* Arousal Tracking (DART) system [210], ethoscope [211], Activity Recording Capillary (ARC) Feeder or CAFE system [212]]. Furthermore, the development of open source software such as Ctrax [213] and JAABA [214], which can be customized to detect various *Drosophila* behaviours, allows video-tracking based methods to be developed to assess additional behaviours including attention [215] and grooming [216].

*Inattention.* Attentional deficit is one of the core features of ADHD, however, attentional processes are also widely affected in other neurodevelopmental and neuropsychiatric disorders, such as intellectual disability (ID), ASD, schizophrenia, and depression [217,218]. In *Drosophila,* attention-like processes have predominantly been studied using vision-based behavioural paradigms (comprehensively reviewed previously [219,220]. Attention-like processes are

classically measured in tethered flight paradigms arena or in Buridan's paradigms (Table 1) [221]. In tethered flight paradigms, a single fly is secured to a torque meter which records changes in flight dynamics in response to visual stimuli presentation [220]. A variation of this paradigm measures walking dynamics of a tethered fly on an air-supported ball [222]. Evidence shows that dopamine levels influence performance in tethered flight paradigms [223,224]. Buridan's paradigms assess fixation strength on visual objects, which is measured by the angle of deviation between the fly's trajectory and either of the two inaccessible visual landmarks [225]. Transient activation of dopamine signalling during development impairs fixation strength in adult [226]. An adapted version of Buridan's paradigm has been used to assess selective attention by measuring behavioural responses of flies to distracting secondary stimuli (Table 1) [215,227]. Classical visual paradigms have been combined with live brain activity recordings through electrophysiology or calcium imaging [228–230]. Such combinatorial approaches allow for more comprehensive insights into attention-like processes in *Drosophila*. Deficits in attention-like processes have been shown in *radish* mutants, a *Drosophila* model of memory consolidation deficits, which could be rescued with methylphenidate [223,230,231]. To date, attention-like behaviours have not been tested in *Drosophila* models of neuropsychiatric disorders.

*Impulsivity.* In other organisms, impulsivity is mostly measured with delay discounting or response inhibition [232]. To date, only few studies have investigated impulsivity in *Drosophila*, mostly assessing impulsivity in form of courtship disinhibition (Table 1). Exposure to psychoactive substances, including ethanol, causes male courtship disinhibition towards both females and other males [233]. Such behaviours were identified to be modulated by dopamine receptors [234].

## 4.2. Established *Drosophila* models of ADHD-related genes

Because of its highly efficient genetics, disease modelling in *Drosophila* so far has mostly focused on characterizing the function of single candidate genes. Several ADHD risk genes have been characterized in recent years that monitored ADHD-related phenotypes in *Drosophila*, particularly increased activity.

*SLC6A3,* also named *DAT,* is one of the earliest identified ADHD-associated genes [235]. As mentioned above, it encodes the dopamine transporter 1 (DAT1) protein. The length of a variable-number tandem repeat (VNTR) at the 3' untranslated region (UTR) of *SLC6A3* correlates with the level of DAT1 [236]. *DAT* mutant flies, also termed *fumin*, exhibit increased dopamine levels and hyperactivity [237,238]. Administering the mood stabilizer valproic acid has been shown to

ameliorate such hyperactivity [239]. In addition, *fumin* flies display deficits in grooming, sleep, and circadian behaviours [238,240,241]. Van der Voet et al. showed that downregulating *Drosophila DAT, Cirl* (the orthologue of *ADGRL3*), and *Nf1* specifically in neurons, respectively, increased activity and reduced sleep, and administering methylphenidate rescued the phenotypes [209]. NF1 (Neurofibromatosis type 1) is a monogenic neurocutaneous syndrome characterized by benign nerve sheath tumors, caused by loss-of-function of the *NF1* gene [242]. Patients with NF1 also suffer from cognitive impairment ranging from learning disabilities to intellectual disability, and are characterized by frequent ADHD features [243–245]. The *Drosophila* model of *Nf1* loss-of-function also displays excessive spontaneous grooming [241]. *ADGRL3* was identified as an ADHD candidate gene from a linkage study based on large multigenerational families in a population isolate [159]. Similar findings have also been reported in zebrafish [145,164] and rodent *ADGRL3* models [165,246], suggesting that the role of *ADGRL3* in ADHD-like behaviours is evolutionary conserved. Another classic ADHD candidate gene is *SLC9A9*, encoding a sodium/proton transporter protein of the solute carrier family. Knockout of *Nhe3*, the *Drosophila SLC9A9* orthologue, caused altered electrophysiology upon visual stimuli, similar to findings in individuals with ASD [247]. The link between *Nhe3* and ADHD-like behaviours in *Drosophila* has yet to be established.

Massive efforts in past decades have identified a multitude of ADHD candidate genes. As most genes are emerging through GWASs, their biological relevance for the etiology of the disorder remains to be demonstrated. Using established genetic tools and behavioural assays in *Drosophila*, it is possible to systematically investigate candidate genes in a high-throughput manner [248]. Among genes linked to the 12 loci associated with ADHD in the latest GWAS meta-analysis [4], *FOXP2* has evoked particular interest. *FOXP2* is a transcription factor of the forkhead box family, which has been in the limelight of research ever since rare mutations were found to cause a severe speech disorder [249], sometimes accompanied with mild cognitive impairment [191]. The *Drosophila FoxP* is highly expressed in the nervous system and is required for synaptic development and dendritic morphogenesis [250]. It plays a role in behaviours such as learning, perceptual decision making, social interaction, and locomotor function [250–253]. Many rodent studies have addressed the role of *FoxP2* in neurogenesis and behaviours, as described elsewhere, suggesting an evolutionary conserved function of *FOXP2*. However, the role of *Drosophila FoxP* in attention or (hyper)activity has yet to be described. Another high confidence risk gene identified in the same GWAS is the transcription factor *MEF2C*. *Foxp2* was found to repress *Mef2c* transcription through DNA binding and repressing *Mef2* rescued vocalization and spinogenesis defects of *Foxp2* knockout mice [254]. In *Drosophila*, *Mef2* has been shown to play a role in ADHD-

related behaviours. *Mef2* is required for maintaining normal circadian behaviour [255,256] and neuronal knockdown of *Mef2* causes increased locomotor activity and sleep loss [257]. Recently, Harich et al. reported a Dutch family with ADHD and cooccurring disorders to segregate with a microduplication in 8p23.3, comprising the *FBXO25* gene [258]. They then continued to functionally validate the newly discovered candidate gene by demonstrating that overexpression of *Drosophila FBOX25* caused ADHD-like behaviours.

## 4.3. Using *Drosophila* to investigate comorbid symptoms of ADHD

A striking feature of ADHD clinical manifestation is the frequent co-occurrence with other neuropsychiatric conditions [49]. This makes further behavioural traits that can be studied in *Drosophila* relevant to ADHD research and modelling. These include habituation learning, working memory, sleep, circadian rhythm, addiction and repetitive behaviour, as well as neuromorphological anomalies reported in ADHD and other disorders.

### *Habituation*

With ASD as the most commonly comorbid condition of ADHD, it is relevant to measure habituation learning. Habituation is a simple, evolutionary conserved form of non-associative learning. It is considered a building block of cognition [259]. It is defined as a decrease in response to a repeated or prolonged harmless stimulus, which cannot be explained by sensory or motor fatigue [259]. Habituation is relevant to ADHD, particularly to (in)attention. Inattention has been considered to arise as the failure to filter out irrelevant environmental stimuli [260,261]. Indeed, evidence suggests that slower habituation in children and adults with ADHD to visual stimuli is correlated with inattention symptoms [262,263]. Habituation was for long considered to result from synaptic depression of excitatory neurons. However, emerging evidence, most importantly from *Drosophila*, demonstrated that habituation can result from potentiation of GABAergic inhibition, a finding that can readily explain a considerable amount of historic literature [264]. Psychiatric disorders, especially ASD and schizophrenia, are hypothesized to be an outcome of imbalance in excitatory/inhibitory activity [265,266].

Habituation can be measured in *Drosophila*, among other assays [267–270], in the light-off jump habituation paradigm (Table 1). This behaviour fulfils all habituation criteria [259,271]. A semi-automated version of the paradigm allows habituation assessment in a high-throughput manner [272]. Using this paradigm, *Drosophila* models of more than a hundred monogenic neurodevelopmental disorders have been shown to display habituation deficits [272–275]. These

include the fly models of the disorders caused by mutations in *NF1* and numerous other genes operating in the Ras-MAPK pathway, as well as genes with synaptic function such as *NRXN1*, *DLG2/3* and *SHANK2/3*, several of these also associated with ADHD-like symptoms in patients. Interestingly, habituation deficits were enriched among those *Drosophila* models that are also associated with ASD [272].

*Working memory*

Working memory has been poorly studied so far in *Drosophila* but of obvious interest given its implication in ADHD and other neuropsychiatric disorders [276]. Meta-analysis studies have shown that individuals with ADHD exhibit verbal and visuo-spatial working memory deficits [277,278]. *Drosophila* has been shown to form visuospatial working memory for objects similar to vertebrates. In 2008, the Strauss lab described a detour setup, in which they showed that flies can remember the position of an object in an arena for several seconds after it has been removed from their environment [279] (Table 1). Strikingly, the very first mutant they identified to display deficits in this paradigm was *ignorant*, the orthologue of the *RPS6KA3* alias *RSK2* gene implicated in Coffin-Lowry syndrome, a severe ID syndrome. This visual and spatial working memory is an attractive paradigm, even more so since its neuronal substrates have been mapped [280] and nitric oxide signalling, a risk pathway for psychiatric disorders [281], has also been implicated in this form of working memory [282]. A recent study started to investigate free-movement patterns in a Y-maze as a measure for spatial working memory and executive function in humans, mice, zebrafish, and fruit flies (Table 1) [283]. They found that flies, like vertebrates, systematically explored the maze, apparently remembering their past positions or choices. Translational efforts across species are of major importance for future perspectives of modelling neurodevelopmental and psychiatric disorders.

*Sleep*

Sleep disturbances are another prominent feature in ADHD. Approximately 25%-50% of children with ADHD report sleep problems [284,285]. Sleep disturbances are also more prevalent in children with neurodevelopmental disorders than in typically developing children [286]. A recent study of Norwegian children revealed that shorter sleep duration was able to predict later psychiatric symptoms [287]. Improving sleep behaviour in child with different neurodevelopmental disorders, including ADHD, has been shown to improve cognition, mood, and behaviours [288].

*Drosophila* is a suitable model to elucidate the role of ADHD genes in sleep. *Drosophila* displays a sleep-like state which possesses key features characterizing sleep: a species-specific posture and/or resting place, modulation by a circadian clock, increased arousal threshold, and a homeostatic response to sleep deprivation [289,290]. Furthermore, similar neurobiological processes are involved in sleep regulation in mammals and in *Drosophila* (reviewed in [291]). Sleep in *Drosophila* can be measured by assessing locomotor activity, as described above. It is defined as a minimum of 5 minutes of inactivity, as arousal threshold significantly increases after 5 minutes of inactivity [292]. Both locomotor activity and sleep have been used to characterize *Drosophila* model of ADHD [209]. Neuronal knockdown of *DAT*, *Cirl*, and *Nf1* have been shown to cause sleep loss, as mentioned above. Also, sleep disturbances have been previously reported in Nrx-1 mutants *Drosophila* (ortholog of *NRXN1*)[293,294] and upon *DISC1* overexpression in a *Drosophila* model of schizophrenia [295].

### *Circadian rhythm*

Defects in circadian rhythm has extensively been connected to many neurodevelopmental and psychiatric disorders, and is increasingly recognized to contribute to the etiology of the disorders rather than only represent a consequence [296,297]. ADHD-associated genes such as *DAT*, *Mef2*, and *period* have been shown to play a role in *Drosophila* circadian rhythm [238,255,256,298,299]. Circadian rhythm is generated by a highly conserved molecular clock, which oscillate in a ~24 hours following the earth rotation period and synchronize behaviours to the time of the day. This molecular clock can be synchronized by environmental cues such as light and temperature. *Drosophila* has been instrumental in understanding these processes at the genetic, biochemical and circuit level, as reviewed in detail elsewhere [300]. Circadian rhythm in *Drosophila* was initially measured by the rate of flies emerging from their pupal case [301,302]. Such method has progressively been replaced by monitoring locomotor activity [303], as already described above.

### *Substance abuse*

Reward and addiction-like behaviours have extensively been studied in *Drosophila*. Assays to test these behaviours, the underlying neuronal circuits and molecular pathways as well as their parallels to the human behaviours have recently been comprehensively reviewed elsewhere [304] (Table 1). Interestingly, a number of circadian genes including *period*, *clock*, *cycle*, and *discs overgrown* have been implicated in ethanol and/or cocaine sensitivity. The classic learning and memory genes *rutaba* and *dunce* have been shown to regulate appetitive memory and ethanol preference [304].

*Repetitive behaviour: Grooming*

Grooming is an evolutionary conserved innate animal behaviour that consists of stereotypical sequences of actions. Grooming is a repetitive behaviour, thus it may be relevant to several neurodevelopmental disorders [305,306]. Fruit flies clean their body parts of dusts and microbes using their legs in a fixed repertoire of cleaning movements [307]. Assessment of fly grooming behaviour can be done by scoring grooming events [216,241] or efficiency, by quantifying the amount of dust removed during grooming [307–309] (Table 1). Several genes have been reported to play a role in grooming behaviour in *Drosophila*, including the D1-like dopamine receptor 1 dDA1 [310]. In addition, abnormal grooming behaviour is observed in *Drosophila* model of Fragile X syndrome [311] and Neurofibromatosis type 1 [241], both characterized by high frequency of ADHD and ASD.

*Neuronal development*

A significant number of human genes associated with neurodevelopmental disorders, including ADHD and its frequent comorbidities, are related to synaptic development and function [312–314]. Post-mortem brain inspection of individuals with ID, ASD, and schizophrenia revealed reduced synaptic counts and dendritic spines [315–317]. Multiple well-established ADHD candidate genes including *DAT1*, *ADGRL3*, *CDH13*, and *SNP25* have been shown to play a role in synaptic development and transmission [318–321]. Additionally, pathway analyses on ADHD GWAS datasets yielded multiple significantly enriched biological pathways related to synaptic function such as transmembrane transport, ion channel activity, and excitatory synapse [4,322,323]. The *Drosophila* neuromuscular junction (NMJ) is a well-established model system to study processes related to synapse formation and synaptic functioning. The *Drosophila* NMJ is glutamatergic and shares major features with excitatory synapses in the mammalian central nervous system [324,325]. Many genes and processes involved in synaptic biology are conserved between *Drosophila* and humans. The NMJ has been used to characterize other *Drosophila* models of neurodevelopmental disorders including ID, ASD, and schizophrenia [197,199,208,326,327].

Changes in dendritic morphology have been reported in multiple neurodevelopmental disorders [328,329]. ADHD genetic risk factors are enriched in genes related to neurite outgrowth [322,330]. Alterations in this process might manifest in later axonal and dendritic anomalies. Thus, neuronal morphology is a relevant phenotype to characterize ADHD candidate genes. In *Drosophila*, the dendritic arborization (da) neurons are a popular model choice to study dendritic morphology. Da neurons are a part of *Drosophila*'s peripheral nervous system; they function as sensory neurons

and show a complex branching pattern [331,332]. Downregulating the expression of *Drosophila* *FOXP2* and *STXBP1* orthologues, in humans associated with various neurodevelopmental disorders [191,249,333,334], has been shown to cause dendritic branching defects [250,335]. In addition, *FOXP2* was recently identified as high confidence ADHD risk gene [4] and mutations in *FOXP2* cause a severe speech disorder with mild cognitive impairment . Mutations in *STXBP1* cause a spectrum of neurodevelopmental disorders, including ID, ASD, ADHD, and epilepsy.

## 4.4. Application of *Drosophila* to investigate therapy options and drug response in ADHD

*Drosophila* is an ideal model to discover novel treatment approaches through unbiased large-scale screening. Many large-scale drug screen studies in *Drosophila* have successfully identified new therapeutic compounds [336,337]. In contrast to classical *in vitro* drug screens, screening in *Drosophila* allows using ADHD-relevant behaviours as readouts, which may increase the chances to discover compounds.

Methylphenidate is one of the most prescribed drugs for treating ADHD symptoms [338]. In flies, methylphenidate has also been shown to ameliorate deficits in attention-like processes in *Drosophila* memory consolidation mutant [230] and hyperactivity-like behaviour in *Drosophila* models of ADHD [209], as already mentioned. Recently, Rohde et al. analysed the genetics underlying the behavioural response to methylphenidate using the Drosophila Genetic Reference Panel (DGRP) [339], a collection of fully sequenced inbred lines derived from a natural population facilitating genotype-phenotype mapping [340]. They identified several genes contributing to variability in the drug response and found that the most active wild-type genotypes became less active upon acute methylphenidate supplementation. These findings argue that the inverted-U shape dose response of methylphenidate is evolutionarily conserved.

It has been proposed that environmental exposure to toxins such as bisphenol A (BPA) contribute to ADHD, particularly in boys [341]. BPA exposure has also been linked to various health issues in humans and animals, including fruit fly [342–345]. Early BPA exposure is associated with increased neuropsychiatric disorders symptoms in children [346–350]. In flies, Kaur et al. reported that BPA exposure to wild-type flies caused abnormal social interaction, reduced locomotion, and increased grooming episodes [343]. Together, these findings illustrate the potential of *Drosophila* to study the effect of exposure to risk-conferring environmental toxin to behaviours.

Additionally, *Drosophila* has recently been shown to be of utility for studying non-pharmacological therapies. In an innovative study, Belfer and colleagues reported that a behavioural regime, resembling sleep opportunity restriction therapy, increases total sleep in short-sleeping *Drosophila* mutants, including DAT1-deficient flies [351]. Sleep restriction therapy (SRT) is widely used as a part of cognitive behaviour therapy for insomnia (CBT-I) and performing SRT alone is sufficient to confer most CBT-I benefits [352]. It is yet to be reported whether SRT may alleviate ADHD-like behavioural alterations. Findings of Belfer and colleagues open a new possibility to use *Drosophila* as a model to investigate the effect of behavioural therapy to disease-relevant phenotypes.

## 4.5. Concluding remarks - *Drosophila*

*Drosophila* is a leading model organism that already provided major breakthroughs in monogenic neurodevelopmental disorders, including the first drug reversal in Fragile X Syndrome [353], the first large-scale approaches to ID/ASD disorders [272,354,355], and countless mechanistic insights into specific genetic disorders. Regardless, the study of ADHD in *Drosophila* is still in its infancy. So far, mostly face-valid behaviours, above all locomotor activity, have been used to investigate specific aspects of ADHD. Here we summarized the already considerable contribution to our understanding of genetics and neurobiology of the disorder. These achievements, with time, are increasing the confidence in the relevance of the applied behaviours and paradigms. Beyond what we discussed above, there likely are many more phenotypical readouts that are relevant to ADHD. Arousal thresholds, for example, is a recurring theme in neurodevelopmental disorders, including ADHD [356,357]. Recently, somatic comorbidities, such as obesity or susceptibility to infection, are moving into the limelight and could be investigated in *Drosophila* models of ADHD [358–364]. Supported by exceptional toolboxes and resources, it is possible to efficiently address the function of candidate genes [365–368], dissect the underlying circuits and mechanisms, and critical time frames during development, with important implications for potential reversibility of the observed defects. In particular, the cost- and time-efficiency of *Drosophila*, together with the high-throughput manner in which some of the above mentioned assays can be conducted makes the fly an organism of choice for approaches highly needed in the ADHD field. These include investigating the large amount of emerging candidate genes and variants that require biological support [369–371], such as the top 50 or 100 findings of the recent GWAS [4]. Testing larger collections of drugs using behavioural assays of confirmed relevance may also generate new breakthroughs [337,372]. Many assays however do need significant set-ups and expertise, which at present may still limit their widespread application.

# REFERENCES

1. Sobanski, E. Psychiatric comorbidity in adults with attention-deficit/hyperactivity disorder (ADHD). *Eur. Arch. Psychiatry Clin. Neurosci.* 2006, *256*, i26–i31.

2. Franke, B.; Faraone, S. V.; Asherson, P.; Buitelaar, J.; Bau, C.H.D.; Ramos-Quiroga, J.A.; Mick, E.; Grevet, E.H.; Johansson, S.; Haavik, J.; et al. The genetics of attention deficit/hyperactivity disorder in adults, a review. *Mol. Psychiatry* 2012, *17*, 960–987.

3. Anttila, V.; Bulik-Sullivan, B.; Finucane, H.K.; Walters, R.K.; Bras, J.; Duncan, L.; Escott-Price, V.; Falcone, G.J.; Gormley, P.; Malik, R.; et al. Analysis of shared heritability in common disorders of the brain. *Science (80-. ).* **2018**, *360*, doi:10.1126/science.aap8757.

4. Demontis, D.; Walters, R.K.; Martin, J.; Mattheisen, M.; Als, T.D.; Agerbo, E.; Baldursson, G.; Belliveau, R.; Bybjerg-Grauholm, J.; Bækvad-Hansen, M.; et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **2019**, *51*, 63–75, doi:10.1038/s41588-018-0269-7.

5. Lee, S.H.; Ripke, S.; Neale, B.M.; Faraone, S. V.; Purcell, S.M.; Perlis, R.H.; Mowry, B.J.; Thapar, A.; Goddard, M.E.; Witte, J.S.; et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **2013**, doi:10.1038/ng.2711.

6. Lee, P.H.; Anttila, V.; Won, H.; Feng, Y.C.A.; Rosenthal, J.; Zhu, Z.; Tucker-Drob, E.M.; Nivard, M.G.; Grotzinger, A.D.; Posthuma, D.; et al. Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell* **2019**, *179*, 1469-1482.e11, doi:10.1016/j.cell.2019.11.020.

7. Gururajan, A.; Reif, A.; Cryan, J.F.; Slattery, D.A. The future of rodent models in depression research. *Nat. Rev. Neurosci.* 2019, *20*, 686–701.

8. Meek, S.; Mashimo, T.; Burdon, T. From engineering to editing the rat genome. *Mamm. Genome* **2017**, *28*, 302–314, doi:10.1007/s00335-017-9705-8.

9. Insel, T.; Cuthbert, B.; Garvey, M.; Heinssen, R.; Pine, D.S.; Quinn, K.; Sanislow, C.; Wang, P. Research Domain Criteria (RDoC): Toward a new classification framework for research on mental disorders. *Am. J. Psychiatry* 2010, *167*, 748–751.

10. Ineichen, C.; Sigrist, H.; Spinelli, S.; Lesch, K.P.; Sautter, E.; Seifritz, E.; Pryce, C.R. Establishing a probabilistic reversal learning test in mice: Evidence for the processes mediating reward-stay and punishment-shift behaviour and for their modulation by serotonin. *Neuropharmacology* **2012**, *63*, 1012–1021, doi:10.1016/j.neuropharm.2012.07.025.

11. Stuart, S.A.; Wood, C.M.; Robinson, E.S.J. Using the affective bias test to predict drug-induced negative affect: implications for drug safety. *Br. J. Pharmacol.* **2017**, *174*, 3200–3210, doi:10.1111/bph.13972.

12. Heisler, J.M.; Morales, J.; Donegan, J.J.; Jett, J.D.; Redus, L.; O'connor, J.C. The attentional set shifting task: A measure of cognitive flexibility in mice. *J. Vis. Exp.* **2015**, e51944, doi:10.3791/51944.

13. Robbins, T.W. The 5-choice serial reaction time task: Behavioural pharmacology and functional neurochemistry. *Psychopharmacology (Berl).* 2002, *163*, 362–380.

14. Kim, C.H.; Hvoslef-Eide, M.; Nilsson, S.R.O.; Johnson, M.R.; Herbert, B.R.; Robbins, T.W.; Saksida, L.M.; Bussey, T.J.; Mar, A.C. The continuous performance test (rCPT) for mice: A novel operant touchscreen test of attentional function. *Psychopharmacology (Berl).* **2015**, *232*, 3947–3966, doi:10.1007/s00213-015-4081-0.

15. Sullivan, J.A.; Dumont, J.R.; Memar, S.; Skirzewski, M.; Wan, J.; Mofrad, M.H.; Ansari, H.Z.; Li, Y.; Muller, L.; Prado, V.F.; et al. New frontiers in translational research: Touchscreens, open science, and the mouse translational research accelerator platform. *Genes. Brain. Behav.* **2021**, *20*, e12705, doi:10.1111/gbb.12705.

16. Mar, A.C.; Robbins, T.W. Delay discounting and impulsive choice in the rat. *Curr. Protoc. Neurosci.* **2007**, *Chapter 8*,

doi:10.1002/0471142301.ns0822s39.

17. Mitchell, S.H. Assessing delay discounting in mice. *Curr. Protoc. Neurosci.* **2014**, *66*, doi:10.1002/0471142301.ns0830s66.

18. Slattery, D.A.; Cryan, J.F. Using the rat forced swim test to assess antidepressant-like activity in rodents. *Nat. Protoc.* 2012, *7*, 1009–1014.

19. Slattery, D.A.; Cryan, J.F. Modelling depression in animals: at the interface of reward and stress pathways. *Psychopharmacology (Berl).* 2017, *234*, 1451–1465.

20. Freudenberg, F.; O'Leary, A.; Aguiar, D.C.; Slattery, D.A. Challenges with modelling anxiety disorders: a possible hindrance for drug discovery. *Expert Opin. Drug Discov.* 2018, *13*, 279–281.

21. Hess, E.; Jinnah, H.; Kozak, C.; Wilson, M. Spontaneous locomotor hyperactivity in a mouse mutant with a deletion including the Snap gene on chromosome 2. *J. Neurosci.* **1992**, *12*, 2865, doi:10.1523/JNEUROSCI.12-07-02865.1992.

22. Sagvolden, T.; Metzger, M.A.; Schiorbeck, H.K.; Rugland, A.L.; Spinnangr, I.; Sagvolden, G. The spontaneously hypertensive rat (SHR) as an animal model of childhood hyperactivity (ADHD): changed reactivity to reinforcers and to psychomotor stimulants. *Behav. Neural Biol.* **1992**, *58*, 103–112, doi:10.1016/0163-1047(92)90315-U.

23. de la Peña, J.B.; dela Peña, I.J.; Custodio, R.J.; Botanas, C.J.; Kim, H.J.; Cheong, J.H. Exploring the Validity of Proposed Transgenic Animal Models of Attention-Deficit Hyperactivity Disorder (ADHD). *Mol. Neurobiol.* 2018, *55*, 3739–3754.

24. Antonucci, F.; Corradini, I.; Fossati, G.; Tomasoni, R.; Menna, E.; Matteoli, M. SNAP-25, a Known presynaptic protein with emerging postsynaptic functions. *Front. Synaptic Neurosci.* **2016**, *8*, 7, doi:10.3389/fnsyn.2016.00007.

25. Corradini, I.; Verderio, C.; Sala, M.; Wilson, M.C.; Matteoli, M. SNAP-25 in neuropsychiatric disorders. In Proceedings of the Annals of the New York Academy of Sciences; Ann N Y Acad Sci, 2009; Vol. 1152, pp. 93–99.

26. Yamada, K. Biological and Pharmaceutical Bulletin: Foreword. *Biol. Pharm. Bull.* 2011, *34*, 1357.

27. Sagvolden, T. Behavioral validation of the spontaneously hypertensive rat (SHR) as an animal model of attention-deficit/hyperactivity disorder (AD/HD). In Proceedings of the Neuroscience and Biobehavioral Reviews; Neurosci Biobehav Rev, 2000; Vol. 24, pp. 31–39.

28. van den Bergh, F.S.; Bloemarts, E.; Chan, J.S.W.; Groenink, L.; Olivier, B.; Oosting, R.S. Spontaneously hypertensive rats do not predict symptoms of attention-deficit hyperactivity disorder. *Pharmacol. Biochem. Behav.* **2006**, *83*, 380–390, doi:10.1016/j.pbb.2006.02.018.

29. Alsop, B. Reprint of "Problems with spontaneously hypertensive rats (SHR) as a model of attention-deficit/hyperactivity disorder (AD/HD)." *J. Neurosci. Methods* **2007**, *166*, XV–XXI, doi:10.1016/j.jneumeth.2006.12.019.

30. Bouchatta, O.; Manouze, H.; Bouali-benazzouz, R.; Kerekes, N.; Ba-M'hamed, S.; Fossat, P.; Landry, M.; Bennis, M. Neonatal 6-OHDA lesion model in mouse induces Attention-Deficit/ Hyperactivity Disorder (ADHD)-like behaviour. *Sci. Rep.* **2018**, *8*, doi:10.1038/s41598-018-33778-0.

31. Efimova, E. V.; Gainetdinov, R.R.; Budygin, E.A.; Sotnikova, T.D. Dopamine transporter mutant animals: A translational perspective. *J. Neurogenet.* **2016**, *30*, 5–15, doi:10.3109/01677063.2016.1144751.

32. Madras, B.K.; Miller, G.M.; Fischman, A.J. The dopamine transporter and attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **2005**, *57*, 1397–1409, doi:10.1016/j.biopsych.2004.10.011.

33. Volkow, N.D.; Wang, G.J.; Newcorn, J.; Fowler, J.S.; Telang, F.; Solanto, M. V.; Logan, J.; Wong, C.; Ma, Y.; Swanson, J.M.; et al. Brain dopamine transporter levels in treatment and drug naïve adults with ADHD. *Neuroimage* **2007**, *34*, 1182–1190, doi:10.1016/j.neuroimage.2006.10.014.

34. Sakrikar, D.; Mazei-Robison, M.S.; Mergy, M.A.; Richtand, N.W.; Han, Q.; Hamilton, P.J.; Bowton, E.; Galli, A.;

Veenstra-Vanderweele, J.; Gill, M.; et al. Attention deficit/hyperactivity disorder-derived coding variation in the dopamine transporter disrupts microdomain targeting and trafficking regulation. *J. Neurosci.* **2012**, *32*, 5385–5397, doi:10.1523/JNEUROSCI.6033-11.2012.

35.  Spencer, T.J.; Biederman, J.; Faraone, S. V; Madras, B.K.; Ali, A.; Dougherty, D.D.; Batchelder, H.; Clarke, A.; Fischman, A.J. Functional genomics of attention-deficit/hyperactivity disorder (ADHD) risk alleles on dopamine transporter binding in ADHD and Healthy Control Subjects. *Biol Psychiatry* **2013**, *74*, 84–89, doi:10.1016/j.biopsych.2012.11.010.Functional.

36.  Zhuang, X.; Oosting, R.S.; Jones, S.R.; Gainetdinov, R.R.; Miller, G.W.; Caron, M.G.; Hen, R. Hyperactivity and impaired response habituation in hyperdopaminergic mice. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 1982–1987, doi:10.1073/pnas.98.4.1982.

37.  Chen, R.; Tilley, M.R.; Wei, H.; Zhou, F.; Zhou, F.M.; Ching, S.; Quan, N.; Stephens, R.L.; Hill, E.R.; Nottoli, T.; et al. Abolished cocaine reward in mice with a cocaine-insensitive dopamine transporter. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 9333–9338, doi:10.1073/pnas.0600905103.

38.  Napolitano, F.; Bonito-Oliva, A.; Federici, M.; Carta, M.; Errico, F.; Magara, S.; Martella, G.; Nisticò, R.; Centonze, D.; Pisani, A.; et al. Role of aberrant striatal dopamine D1 receptor/cAMP/protein kinase A/DARPP32 signaling in the paradoxical calming effect of amphetamine. *J. Neurosci.* **2010**, *30*, 11043–11056, doi:10.1523/JNEUROSCI.1682-10.2010.

39.  Yan, T.C.; Hunt, S.P.; Stanford, S.C. Behavioural and neurochemical abnormalities in mice lacking functional tachykinin-1 (NK1) receptors: A model of attention deficit hyperactivity disorder. *Neuropharmacology* **2009**, *57*, 627–635, doi:10.1016/j.neuropharm.2009.08.021.

40.  Yan, T.; Mcquillin, A.; Thapar, A.; Asherson, P.; Hunt, S.; Stanford, S.; Gurling, H. NK1 (TACR1) receptor gene 'knockout' mouse phenotype predicts genetic association with ADHD. **2010**, *24*, 27–38, doi:10.1177/0269881108100255.

41.  Siesser, W.B.; Cheng, S.Y.; McDonald, M.P. Hyperactivity, impaired learning on a vigilance task, and a differential response to methylphenidate in the TRβPV knock-in mouse. *Psychopharmacology (Berl).* **2005**, *181*, 653–663, doi:10.1007/s00213-005-0024-5.

42.  Siesser, W.B.; Zhao, J.; Miller, L.R.; Cheng, S.Y.; McDonald, M.P. Transgenic mice expressing a human mutant β1 thyroid receptor are hyperactive, impulsive, and inattentive. *Genes, Brain Behav.* **2006**, *5*, 282–297, doi:10.1111/j.1601-183X.2005.00161.x.

43.  Higgins, G.A.; Breysse, N. Rodent Model of Attention: The 5-Choice Serial Reaction Time Task. *Curr. Protoc. Pharmacol.* **2008**, *41*, 5.49.1-5.49.20, doi:10.1002/0471141755.ph0549s41.

44.  de Leeuw, C.A.; Mooij, J.M.; Heskes, T.; Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput. Biol.* **2015**, *11*, 4, doi:10.1371/journal.pcbi.1004219.

45.  Prince, J. Catecholamine dysfunction in attention-deficit/hyperactivity disorder an update. *J. Clin. Psychopharmacol.* **2008**, *28*, 39–45, doi:10.1097/JCP.0b013e318174f92a.

46.  Del Campo, N.; Chamberlain, S.R.; Sahakian, B.J.; Robbins, T.W. The roles of dopamine and noradrenaline in the pathophysiology and treatment of attention-deficit/hyperactivity disorder. *Biol. Psychiatry* **2011**, *69*, e145–e157, doi:10.1016/j.biopsych.2011.02.036.

47.  Arnsten, A.F.T. Fundamentals of attention-deficit/hyperactivity disorder: Circuits and pathways. *J. Clin. Psychiatry* **2006**, *67*, 7–12.

48.  Mota, N.R.; Poelmans, G.; Klein, M.; Torrico, B.; Fernàndez-Castillo, N.; Cormand, B.; Reif, A.; Franke, B.; Arias Vásquez, A. Cross-disorder genetic analyses implicate dopaminergic signaling as a biological link between Attention-Deficit/Hyperactivity Disorder and obesity measures. *Neuropsychopharmacology* **2020**, *45*, 1188–1195,

doi:10.1038/s41386-019-0592-4.

49. Katzman, M.A.; Bilkey, T.S.; Chokka, P.R.; Fallu, A.; Klassen, L.J. Adult ADHD and comorbid disorders: clinical implications of a dimensional approach. *BMC Psychiatry* **2017**, *17*, 302, doi:10.1186/s12888-017-1463-3.

50. Rommelse, N.N.J.; Franke, B.; Geurts, H.M.; Hartman, C.A.; Buitelaar, J.K. Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur. Child Adolesc. Psychiatry* **2010**, *19*, 281–295, doi:10.1007/s00787-010-0092-x.

51. Rommelse, N.N.J.; Geurts, H.M.; Franke, B.; Buitelaar, J.K.; Hartman, C.A. A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neurosci. Biobehav. Rev.* **2011**, *35*, 1363–1396, doi:10.1016/j.neubiorev.2011.02.015.

52. Lord, C.; Elsabbagh, M.; Baird, G.; Veenstra-Vanderweele, J. Autism spectrum disorder. *Lancet* **2018**, *392*, 508–520, doi:10.1016/S0140-6736(18)31129-2.

53. Ey, E.; Leblond, C.S.; Bourgeron, T. Behavioral profiles of mouse models for autism spectrum disorders. *Autism Res.* **2011**, *4*, 5–16, doi:10.1002/aur.175.

54. Kazdoba, T.M.; Prescott, T.L.; Yang, M.; Silverman, J.L.; Solomon, M.; Crawley, J.N. Translational Mouse Models of Autism: Advancing Toward Pharmacological Therapeutics. *Transl. Neuropsychopharmacol. Curr. Top. Behav. Neurosci.* **2016**, *28*, 1–52, doi:10.1007/7854_2015_5003.

55. Silverman, J.L.; Yang, M.; Lord, C.; Crawley, J.N. Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* **2010**, *11*, 490–502, doi:10.1038/nrn2851.

56. Crawley, J.N. Mouse behavioral assays relevant to the symptoms of autism. *Brain Pathol.* **2007**, *17*, 448–459, doi:10.1111/j.1750-3639.2007.00096.x.

57. Peñagarikano, O.; Abrahams, B.S.; Herman, E.I.; Winden, K.D.; Gdalyahu, A.; Dong, H.; Sonnenblick, L.I.; Gruver, R.; Almajano, J.; Bragin, A.; et al. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **2011**, *147*, 235–246, doi:10.1016/j.cell.2011.08.040.

58. Zang, J.B.; Nosyreva, E.D.; Spencer, C.M.; Volk, L.J.; Musunuru, K.; Zhong, R.; Stone, E.F.; Yuva-Paylor, L.A.; Huber, K.M.; Paylor, R.; et al. A mouse model of the human fragile X syndrome I304N mutation. *PLoS Genet.* **2009**, *5*, doi:10.1371/journal.pgen.1000758.

59. Han, K.; Holder, J.L.; Schaaf, C.P.; Lu, H.-C.; Chen, H.; Kang, H.; Tang, J.; Wu, Z.; Hao, S.; Cheung, S.W.; et al. SHANK3 overexpression causes manic-like behaviour with unique pharmacogenetic properties. *Nature* **2013**, *503*, 72–77, doi:10.1038/nature12630.

60. Won, H.; Lee, H.R.; Gee, H.Y.; Mah, W.; Kim, J.I.; Lee, J.; Ha, S.; Chung, C.; Jung, E.S.; Cho, Y.S.; et al. Autistic-like social behaviour in Shank2-mutant mice improved by restoring NMDA receptor function. *Nature* **2012**, *486*, 261–265, doi:10.1038/nature11208.

61. Eskander, N. The Psychosocial Outcome of Conduct and Oppositional Defiant Disorder in Children With Attention Deficit Hyperactivity Disorder. *Cureus* **2020**, *12*, doi:10.7759/cureus.9521.

62. Young, S.; Thome, J. ADHD and offenders. *World J. Biol. Psychiatry* **2011**, *12*, 124–128, doi:10.3109/15622975.2011.600319.

63. Sebastian, A.; Retz, W.; Tüscher, O.; Turner, D. Violent offending in borderline personality disorder and attention deficit/hyperactivity disorder. *Neuropharmacology* 2019, *156*, 107565.

64. Demontis, D.; Walters, R.; Rajagopal, V.M.; Waldman, I.D.; Grove, J.; Als, T.D.; Dalsgaard, S.; Ribases, M.; Grauholm, J.; Bækvad-Hansen, M.; et al. Identification of risk variants and characterization of the polygenic architecture of disruptive behavior disorders in the context of ADHD. *bioRxiv* **2019**, 1–34, doi:10.1101/791160.

65. Freudenberg, F.; Carreño Gutierrez, H.; Post, A.M.; Reif, A.; Norton, W.H.J. Aggression in non-human

vertebrates: Genetic mechanisms and molecular pathways. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2016**, *171*, 603–640, doi:10.1002/ajmg.b.32358.

66. Takahashi, A.; Miczek, K.A. Neurogenetics of Aggressive Behavior: Studies in Rodents. *Neurosci. Aggress. Curr. Top. Behav. Neurosci.* **2013**, *17*, 3–44, doi:https://doi.org/10.1007/7854_2013_263.

67. Fernàndez-Castillo, N.; Cormand, B. Aggressive behavior in humans: Genes and pathways identified through association studies. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 2016, *171*, 676–696.

68. Pinsonneault, J.K.; Frater, J.T.; Kompa, B.; Mascarenhas, R.; Wang, D.; Sadee, W. Intronic SNP in ESR1 encoding human estrogen receptor alpha is associated with brain ESR1 mRNA isoform expression and behavioral traits. *PLoS One* **2017**, *12*, e0179020, doi:10.1371/journal.pone.0179020.

69. Levy, F. Synaptic gating and ADHD: A biological theory of comorbidity of ADHD and anxiety. *Neuropsychopharmacology* 2004.

70. Schatz, D.B.; Rostain, A.L. ADHD with comorbid anxiety. A review of the current literature. *J. Atten. Disord.* 2006.

71. Himanshu; Dharmila; Sarkar, D.; Nutan A review of behavioral tests to evaluate different types of anxiety and anti-anxiety effects. *Clin. Psychopharmacol. Neurosci.* 2020.

72. Zhang, N.; Zhong, P.; Shin, S.M.; Metallo, J.; Danielson, E.; Olsen, C.M.; Liu, Q.S.; Lee, S.H. S-scam, a rare copy number variation gene, induces schizophrenia-related endophenotypes in transgenic mouse model. *J. Neurosci.* **2015**, *35*, 1892–1904, doi:10.1523/JNEUROSCI.3658-14.2015.

73. Chen, C.; Rainnie, D.G.; Greene, R.W.; Tonegawa, S. Abnormal fear response and aggressive behavior in mutant mice deficient for α-calcium-calmodulin kinase II. *Science (80-. ).* **1994**, *266*, 291–294, doi:10.1126/science.7939668.

74. Lee, P.C.W.; Dodart, J.C.; Aron, L.; Finley, L.W.; Bronson, R.T.; Haigis, M.C.; Yankner, B.A.; Harper, J.W. Altered Social Behavior and Neuronal Development in Mice Lacking the Uba6-Use1 Ubiquitin Transfer System. *Mol. Cell* **2013**, *50*, 172–184, doi:10.1016/j.molcel.2013.02.014.

75. Yamasaki, N.; Maekawa, M.; Kobayashi, K.; Kajii, Y.; Maeda, J.; Soma, M.; Takao, K.; Tanda, K.; Ohira, K.; Toyama, K.; et al. Alpha-CaMKII deficiency causes immature dentate gyrus, a novel candidate endophenotype of psychiatric disorders. *Mol. Brain* **2008**, *1*, 6, doi:10.1186/1756-6606-1-6.

76. Sandau, U.S.; Alderman, Z.; Corfas, G.; Ojeda, S.R.; Raber, J. Astrocyte-specific disruption of SynCAM1 signaling results in ADHD-like behavioral manifestations. *PLoS One* **2012**, *7*, doi:10.1371/journal.pone.0036424.

77. Drapeau, E.; Riad, M.; Kajiwara, Y.; Buxbaum, J.D. Behavioral phenotyping of an improved mouse model of Phelan-McDermid Syndrome with a complete deletion of the Shank3 gene. *bioRxiv* **2018**, *5*, doi:10.1101/278622.

78. Darwish, A.H.; Elgohary, T.M.; Nosair, N.A. Serum Interleukin-6 Level in Children With Attention-Deficit Hyperactivity Disorder (ADHD). *J. Child Neurol.* **2019**, *34*, 61–67, doi:10.1177/0883073818809831.

79. Elhady, M.; Elattar, R.S.; Elaidy, A.M.A.; Abdallah, N.A.; Elmalt, H.A. Role of inflammation in childhood epilepsy and ADHD comorbidity. *Appl. Neuropsychol. Child* **2020**, doi:10.1080/21622965.2020.1807982.

80. Chang, J.P.C.; Mondelli, V.; Satyanarayanan, S.K.; Chiang, Y.J.; Chen, H.T.; Su, K.P.; Pariante, C.M. Cortisol, inflammatory biomarkers and neurotrophins in children and adolescents with attention deficit hyperactivity disorder (ADHD) in Taiwan. *Brain. Behav. Immun.* **2020**, *88*, 105–113, doi:10.1016/j.bbi.2020.05.017.

81. Key, K.V.; Mudd-Martin, G.; Moser, D.K.; Rayens, M.K.; Morford, L.A. Inflammatory Genotype Moderates the Association Between Anxiety and Systemic Inflammation in Adults at Risk for Cardiovascular Disease. *J. Cardiovasc. Nurs.* **2020**, doi:10.1097/jcn.0000000000000742.

82. Renna, M.E.; O'Toole, M.S.; Spaeth, P.E.; Lekander, M.; Mennin, D.S. The association between anxiety, traumatic stress, and obsessive–compulsive disorders and chronic inflammation: A systematic review and meta-analysis. *Depress. Anxiety* **2018**, doi:10.1002/da.22790.

83. English, B.A.; Hahn, M.K.; Gizer, I.R.; Mazei-Robison, M.; Steele, A.; Kurnik, D.M.; Stein, M.A.; Waldman, I.D.;

Blakely, R.D. Choline transporter gene variation is associated with attention-deficit hyperactivity disorder. *J. Neurodev. Disord.* **2009**, *1*, 252–263, doi:10.1007/s11689-009-9033-8.

84. Bron, T.I.; Bijlenga, D.; Verduijn, J.; Penninx, B.W.J.H.; Beekman, A.T.F.; Kooij, J.J.S. Prevalence of ADHD symptoms across clinical stages of major depressive disorder. *J. Affect. Disord.* **2016**, *197*, 29–35, doi:10.1016/j.jad.2016.02.053.

85. Du Rietz, E.; Coleman, J.; Glanville, K.; Choi, S.W.; O'Reilly, P.F.; Kuntsi, J. Association of Polygenic Risk for Attention-Deficit/Hyperactivity Disorder With Co-occurring Traits and Disorders. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2018**, *3*, 635–643, doi:10.1016/j.bpsc.2017.11.013.

86. Agudelo, L.Z.; Femenía, T.; Orhan, F.; Porsmyr-Palmertz, M.; Goiny, M.; Martinez-Redondo, V.; Correia, J.C.; Izadi, M.; Bhat, M.; Schuppe-Koistinen, I.; et al. Skeletal muscle PGC-1α1 modulates kynurenine metabolism and mediates resilience to stress-induced depression. *Cell* **2014**, *159*, 33–45, doi:10.1016/j.cell.2014.07.051.

87. Lin, J.; Wu, P.H.; Tarr, P.T.; Lindenberg, K.S.; St-Pierre, J.; Zhang, C.Y.; Mootha, V.K.; Jäger, S.; Vianna, C.R.; Reznick, R.M.; et al. Defects in adaptive energy metabolism with CNS-linked hyperactivity in PGC-1α null mice. *Cell* **2004**, *119*, 121–135, doi:10.1016/j.cell.2004.09.013.

88. Butterweck, V.; Prinz, S.; Schwaninger, M. The role of interleukin-6 in stress-induced hyperthermia and emotional behaviour in mice. *Behav. Brain Res.* **2003**, *144*, 49–56, doi:10.1016/S0166-4328(03)00059-7.

89. Yankelevitch-Yahav, R.; Franko, M.; Huly, A.; Doron, R. The forced swim test as a model of depressive-like behavior. *J. Vis. Exp.* **2015**, *2015*, doi:10.3791/52587.

90. Rojas, P.; Joodmardi, E.; Hong, Y.; Perlmann, T.; Ögren, S.O. Adult mice with reduced Nurr1 expression: An animal model for schizophrenia. *Mol. Psychiatry* **2007**, *12*, 756–766, doi:10.1038/sj.mp.4001993.

91. Ferguson, G.D.; Anagnostaras, S.G.; Silva, A.J.; Herschman, H.R. Deficits in memory and motor performance in synaptotagmin IV mutant mice. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5598–5603, doi:10.1073/pnas.100104597.

92. Chourbaji, S.; Urani, A.; Inta, I.; Sanchis-Segura, C.; Brandwein, C.; Zink, M.; Schwaninger, M.; Gass, P. IL-6 knockout mice exhibit resistance to stress-induced development of depression-like behaviors. *Neurobiol. Dis.* **2006**, *23*, 587–594, doi:10.1016/j.nbd.2006.05.001.

93. Donev, R.; Gantert, D.; Alawam, K.; Edworthy, A.; Häßler, F.; Meyer-Lindenberg, A.; Dressing, H.; Thome, J. Comorbidity of schizophrenia and adult attention-deficit hyperactivity disorder. *World J. Biol. Psychiatry* **2011**, *12*, 52–56, doi:10.3109/15622975.2011.599212.

94. Amann, L.C.; Gandal, M.J.; Halene, T.B.; Ehrlichman, R.S.; White, S.L.; McCarren, H.S.; Siegel, S.J. Mouse behavioral endophenotypes for schizophrenia. *Brain Res. Bull.* **2010**, *83*, 147–161, doi:10.1016/j.brainresbull.2010.04.008.

95. Powell, S.B. Genetic Models of Sensorimotor Gating: Relevance to Neuropsychiatric Disorders. *Curr. Top. Behav. Neurosci.* **2012**, *12*, 251–318, doi:10.1007/7854_2011_195.

96. Van Den Buuse, M. Modeling the positive symptoms of schizophrenia in genetically modified mice: Pharmacology and methodology aspects. *Schizophr. Bull.* **2010**, *36*, 246–270, doi:10.1093/schbul/sbp132.

97. Carbonell, A.U.; Cho, C.H.; Tindi, J.O.; Counts, P.A.; Bates, J.C.; Erdjument-Bromage, H.; Cvejic, S.; Iaboni, A.; Kvint, I.; Rosensaft, J.; et al. Haploinsufficiency in the ANKS1B gene encoding AIDA-1 leads to a neurodevelopmental syndrome. *Nat. Commun.* **2019**, *10*, doi:10.1038/s41467-019-11437-w.

98. Weber, H.; Kittel-Schneider, S.; Gessner, A.; Domschke, K.; Neuner, M.; Jacob, C.P.; Buttenschon, H.N.; Boreatti-Hümmer, A.; Volkert, J.; Herterich, S.; et al. Cross-disorder analysis of bipolar risk genes: Further evidence of dGKH as a risk gene for bipolar disorder, but also unipolar depression and adult ADHD. *Neuropsychopharmacology* **2011**, doi:10.1038/npp.2011.98.

99. Berryer, M.H.; Hamdan, F.F.; Klitten, L.L.; Møller, R.S.; Carmant, L.; Schwartzentruber, J.; Patry, L.; Dobrzeniecka,

S.; Rochefort, D.; Neugnot-Cerioli, M.; et al. Mutations in SYNGAP1 Cause Intellectual Disability, Autism, and a Specific Form of Epilepsy by Inducing Haploinsufficiency. *Hum. Mutat.* **2013**, *34*, 385–394, doi:10.1002/humu.22248.

100. Piñeiro-Dieguez, B.; Balanzá-Martínez, V.; García-García, P.; Soler-López, B.; Domingo, M.A.; Labarra, J.D.A.; Lobato, P.A.; Salamanca, A.A.; Bes, J.A.; Fernández, F.J.B.; et al. Psychiatric Comorbidity at the Time of Diagnosis in Adults With ADHD: The CAT Study. *J. Atten. Disord.* **2016**, *20*, 1066–1075, doi:10.1177/1087054713518240.

101. Konstenius, M.; Leifman, A.; van Emmerik-van Oortmerssen, K.; van de Glind, G.; Franck, J.; Moggi, F.; Ramos-Quiroga, J.A.; Levin, F.R.; Carpentier, P.J.; Skutle, A.; et al. Childhood trauma exposure in substance use disorder patients with and without ADHD. *Addict. Behav.* **2017**, *65*, 118–124, doi:10.1016/j.addbeh.2016.10.016.

102. Green, J.G.; McLaughlin, K.A.; Berglund, P.A.; Gruber, M.J.; Sampson, N.A.; Zaslavsky, A.M.; Kessler, R.C. Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication I: Associations with first onset of DSM-IV disorders. *Arch. Gen. Psychiatry* **2010**, *67*, 113–123, doi:10.1001/archgenpsychiatry.2009.186.

103. Treur, J.L.; Demontis, D.; Smith, G.D.; Sallis, H.; Richardson, T.G.; Wiers, R.W.; Børglum, A.D.; Verweij, K.J.H.; Munafò, M.R. Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. *Addict. Biol.* **2019**, doi:10.1111/adb.12849.

104. Martelle, J.L.; Nader, M.A. Animal Models of Drug Addiction: Cocaine. In *Biological Research on Addiction*; Academic Press, 2013; pp. 79–87 ISBN 9780123983350.

105. Gerber, D.J.; Sotnikova, T.D.; Gainetdinov, R.R.; Huang, S.Y.; Caron, M.G.; Tonegawa, S. Hyperactivity, elevated dopaminergic transmission, and response to amphetamine in M1 muscarinic acetylcholine receptor-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 15312–15317, doi:10.1073/pnas.261583798.

106. Bello, E.P.; Mateo, Y.; Gelman, D.M.; Noaín, D.; Shin, J.H.; Low, M.J.; Alvarez, V.A.; Lovinger, D.M.; Rubinstein, M. Cocaine supersensitivity and enhanced motivation for reward in mice lacking dopamine D2 autoreceptors. *Nat. Neurosci.* **2011**, *14*, 1033–1038, doi:10.1038/nn.2862.

107. Vekovischeva, O.Y.; Zamanillo, D.; Echenko, O.; Seppälä, T.; Uusi-Oukari, M.; Honkanen, A.; Seeburg, P.H.; Sprengel, R.; Korpi, E.R. Morphine-induced dependence and sensitization are altered in mice deficient in AMPA-type glutamate receptor-A subunits. *J. Neurosci.* **2001**, *21*, 4451–4459, doi:10.1523/jneurosci.21-12-04451.2001.

108. Vekovischeva, O.Y.; Aitta-Aho, T.; Echenko, O.; Kankaanpää, A.; Seppälä, T.; Honkanen, A.; Sprengel, R.; Korpi, E.R. Reduced aggression in AMPA-type glutamate receptor GluR-A subunit-deficient mice. *Genes, Brain Behav.* **2004**, *3*, 253–265, doi:10.1111/j.1601-1848.2004.00075.x.

109. Kirshenbaum, G.S.; Clapcote, S.J.; Duffy, S.; Burgess, C.R.; Petersen, J.; Jarowek, K.J.; Yücel, Y.H.; Cortez, M.A.; Snead, O.C.; Vilsen, B.; et al. Mania-like behavior induced by genetic dysfunction of the neuron-specific Na+,K+-ATPase $\alpha$3 sodium pump. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 18144–18149, doi:10.1073/pnas.1108416108.

110. Quintana, A.; Sanz, E.; Wang, W.; Storey, G.P.; Güler, A.D.; Wanat, M.J.; Roller, B.A.; La Torre, A.; Amieux, P.S.; McKnight, G.S.; et al. Lack of GPR88 enhances medium spiny neuron activity and alters motor- and cue-dependent behaviors. *Nat. Neurosci.* **2012**, *15*, 1547–1555, doi:10.1038/nn.3239.

111. Lu, H.C.; Tan, Q.; Rousseaux, M.W.C.; Wang, W.; Kim, J.Y.; Richman, R.; Wan, Y.W.; Yeh, S.Y.; Patel, J.M.; Liu, X.; et al. Disruption of the ATXN1-CIC complex causes a spectrum of neurobehavioral phenotypes in mice and humans. *Nat. Genet.* **2017**, *49*, 527–536, doi:10.1038/ng.3808.

112. Xu, M.; Hu, X.T.; Cooper, D.C.; Moratalla, R.; Graybiel, A.M.; White, F.J.; Tonegawa, S. Elimination of cocaine-induced hyperactivity and dopamine-mediated neurophysiological effects in dopamine D1 receptor mutant mice. *Cell* **1994**, *79*, 945–955, doi:10.1016/0092-8674(94)90026-4.

113. Francès, H.; Le Foll, B.; Diaz, J.; Smirnova, M.; Sokoloff, P. Role of DRD3 in morphine-induced conditioned place

preference using drd3-knockout mice. *Neuroreport* **2004**, *15*, 2245–2249, doi:10.1097/00001756-200410050-00021.

114. Acilli, D.; Fishburn, C.S.; Drago, J.; Steiner, H.; Lachowicz, J.E.; Park -, B.H.; Gauda, E.B.; Lee, E.J.; Cool, M.H.; Sibley, D.R.; et al. A targeted mutation of the D3 dopamine receptor gene is associated with hyperactivity in mice. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 1945–1949, doi:10.1073/pnas.93.5.1945.

115. Cox, M.M.; Tucker, A.M.; Tang, J.; Talbot, K.; Richer, D.C.; Yeh, L.; Arnold, S.E. Neurobehavioral abnormalities in the dysbindin-1 mutant, sandy, on a C57BL/6J genetic background. *Genes, Brain Behav.* **2009**, *8*, 390–397, doi:10.1111/j.1601-183X.2009.00477.x.

116. Carr, G. V.; Jenkins, K.A.; Weinberger, D.R.; Papaleo, F. Loss of dysbindin-1 in mice impairs reward-based operant learning by increasing impulsive and compulsive behavior. *Behav. Brain Res.* **2013**, *241*, 173–184, doi:10.1016/j.bbr.2012.12.021.

117. Gudmundsson, O.O.; Walters, G.B.; Ingason, A.; Johansson, S.; Zayats, T.; Athanasiu, L.; Sonderby, I.E.; Gustafsson, O.; Nawaz, M.S.; Jonsson, G.F.; et al. Attention-deficit hyperactivity disorder shares copy number variant risk with schizophrenia and autism spectrum disorder. *Transl. Psychiatry* **2019**, *9*, doi:10.1038/s41398-019-0599-y.

118. Portmann, T.; Yang, M.; Mao, R.; Panagiotakos, G.; Ellegood, J.; Dolen, G.; Bader, P.L.; Grueter, B.A.; Goold, C.; Fisher, E.; et al. Behavioral abnormalities and circuit defects in the basal ganglia of a mouse model of 16p11.2 deletion syndrome. *Cell Rep.* **2014**, *7*, 1077–1092, doi:10.1016/j.celrep.2014.03.036.

119. Horev, G.; Ellegood, J.; Lerch, J.P.; Son, Y.-E.E.; Muthuswamy, L.; Vogel, H.; Krieger, A.M.; Buja, A.; Henkelman, R.M.; Wigler, M.; et al. Dosage-dependent phenotypes in models of 16p11.2 lesions found in autism. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17076–81, doi:10.1073/pnas.1114042108.

120. Gnanavel, S. Smith-Magenis syndrome: Behavioural phenotype mimics ADHD. *BMJ Case Rep.* **2014**, *2014*, doi:10.1136/bcr-2013-201766.

121. Walz, K.; Spencer, C.; Kaasik, K.; Lee, C.C.; Lupski, J.R.; Paylor, R. Behavioral characterization of mouse models for Smith-Magenis syndrome and dup(17)(p11.2p11.2). *Hum. Mol. Genet.* **2004**, *13*, 367–378, doi:10.1093/hmg/ddh044.

122. Homberg, J.R.; Kyzar, E.J.; Nguyen, M.; Norton, W.H.; Pittman, J.; Poudel, M.K.; Gaikwad, S.; Nakamura, S.; Koshiba, M.; Yamanouchi, H.; et al. Understanding autism and other neurodevelopmental disorders through experimental translational neurobehavioral models. *Neurosci. Biobehav. Rev.* 2016, *65*, 292–312.

123. Adinolfi, A.; Zelli, S.; Leo, D.; Carbone, C.; Mus, L.; Illiano, P.; Alleva, E.; Gainetdinov, R.R.; Adriani, W. Behavioral characterization of DAT-KO rats and evidence of asocial-like phenotypes in DAT-HET rats: The potential involvement of norepinephrine system. *Behav. Brain Res.* **2019**, *359*, 516–527, doi:10.1016/j.bbr.2018.11.028.

124. Mortimer, N.; Ganster, T.; O'Leary, A.; Popp, S.; Freudenberg, F.; Reif, A.; Soler Artigas, M.; Ribasés, M.; Ramos-Quiroga, J.A.; Lesch, K.P.; et al. Dissociation of impulsivity and aggression in mice deficient for the ADHD risk gene Adgrl3: Evidence for dopamine transporter dysregulation. *Neuropharmacology* **2019**, *156*, doi:10.1016/j.neuropharm.2019.02.039.

125. Regan, S.L.; Hufgard, J.R.; Pitzer, E.M.; Sugimoto, C.; Hu, Y.C.; Williams, M.T.; Vorhees, C. V. Knockout of latrophilin-3 in Sprague-Dawley rats causes hyperactivity, hyper-reactivity, under-response to amphetamine, and disrupted dopamine markers. *Neurobiol. Dis.* **2019**, *130*, doi:10.1016/j.nbd.2019.104494.

126. Jurgensen, S.; Castillo, P.E. Selective dysregulation of hippocampal inhibition in the mouse lacking autism candidate gene CNTNAP2. *J. Neurosci.* **2015**, *35*, 14681–14687, doi:10.1523/JNEUROSCI.1666-15.2015.

127. Vecchia, E.D.; Mortimer, N.; Palladino, V.S.; Kittel-Schneider, S.; Lesch, K.P.; Reif, A.; Schenck, A.; Norton, W.H.J. Cross-species models of attention-deficit/hyperactivity disorder and autism spectrum disorder: Lessons from

CNTNAP2, ADGRL3, and PARK2. *Psychiatr. Genet.* **2019**, *29*, 1–17, doi:10.1097/YPG.0000000000000211.

128. Sanjuán, J.; Castro-Martínez, X.H.; García-Martí, G.; González-Fernández, J.; Sanz-Requena, R.; Haro, J.M.; Meana, J.J.; Martí-Bonmatí, L.; Nacher, J.; Sebastiá-Ortega, N.; et al. FOXP2 expression and gray matter density in the male brains of patients with schizophrenia. *Brain Imaging Behav.* **2020**, doi:10.1007/s11682-020-00339-x.

129. Tolosa, A.; Sanjuán, J.; Dagnall, A.M.; Moltó, M.D.; Herrero, N.; de Frutos, R. FOXP2 gene and language impairment in schizophrenia: Association and epigenetic studies. *BMC Med. Genet.* **2010**, *11*, doi:10.1186/1471-2350-11-114.

130. Oswald, F.; Klöble, P.; Ruland, A.; Rosenkranz, D.; Hinz, B.; Butter, F.; Ramljak, S.; Zechner, U.; Herlyn, H. The FOXP2-Driven network in developmental disorders and neurodegeneration. *Front. Cell. Neurosci.* **2017**, *11*, doi:10.3389/fncel.2017.00212.

131. Co, M.; Hickey, S.L.; Kulkarni, A.; Harper, M.; Konopka, G. Cortical Foxp2 Supports Behavioral Flexibility and Developmental Dopamine D1 Receptor Expression. *Cereb. Cortex* **2020**, *30*, 1855–1870, doi:10.1093/cercor/bhz209.

132. Medvedeva, V.P.; Rieger, M.A.; Vieth, B.; Mombereau, C.; Ziegenhain, C.; Ghosh, T.; Cressant, A.; Enard, W.; Granon, S.; Dougherty, J.D.; et al. Altered social behavior in mice carrying a cortical Foxp2 deletion. *Hum. Mol. Genet.* **2019**, *28*, 701–717, doi:10.1093/hmg/ddy372.

133. Muench, C.; Schwandt, M.; Jung, J.; Cortes, C.R.; Momenan, R.; Lohoff, F.W. The major depressive disorder GWAS-supported variant rs10514299 in TMEM161B-MEF2C predicts putamen activation during reward processing in alcohol dependence. *Transl. Psychiatry* **2018**, *8*, doi:10.1038/s41398-018-0184-9.

134. Ruan, J.W.; Statt, S.; Huang, C.T.; Tsai, Y.T.; Kuo, C.C.; Chan, H.L.; Liao, Y.C.; Tan, T.H.; Kao, C.Y. Dual-specificity phosphatase 6 deficiency regulates gut microbiome and transcriptome response against diet-induced obesity in mice. *Nat. Microbiol.* **2016**, *2*, doi:10.1038/nmicrobiol.2016.220.

135. Curado, S.; Anderson, R.M.; Jungblut, B.; Mumm, J.; Schroeter, E.; Stainier, D.Y.R. Conditional targeted cell ablation in zebrafish: A new tool for regeneration studies. *Dev. Dyn.* **2007**, *236*, 1025–1035, doi:10.1002/dvdy.21100.

136. Albadri, S.; De Santis, F.; Di Donato, V.; Del Bene, F. CRISPR/Cas9-mediated knockin and knockout in Zebrafish. In *Research and Perspectives in Neurosciences*; Springer Verlag, 2017; pp. 41–49.

137. Förster, D.; Kramer, A.; Baier, H.; Kubo, F. Optogenetic precision toolkit to reveal form, function and connectivity of single neurons. *Methods* **2018**, *150*, 42–48, doi:10.1016/j.ymeth.2018.08.012.

138. Norton, W. Towards developmental models of psychiatric disorders in zebrafish. *Front. Neural Circuits* **2013**, *7*, doi:10.3389/fncir.2013.00079.

139. Panula, P.; Chen, Y.C.; Priyadarshini, M.; Kudo, H.; Semenova, S.; Sundvik, M.; Sallinen, V. The comparative neuroanatomy and neurochemistry of zebrafish CNS systems of relevance to human neuropsychiatric diseases. *Neurobiol. Dis.* **2010**, *40*, 46–57, doi:10.1016/j.nbd.2010.05.010.

140. Rihel, J.; Prober, D.A.; Arvanites, A.; Lam, K.; Zimmerman, S.; Jang, S.; Haggarty, S.J.; Kokel, D.; Rubin, L.L.; Peterson, R.T.; et al. Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science (80-. ).* **2010**, *327*, 348–351, doi:10.1126/science.1183090.

141. Norton, W.H.J. Measuring Larval Zebrafish Behavior: Locomotion, Thigmotaxis, and Startle. In; Humana Press, Totowa, NJ, 2012; pp. 3–20.

142. Jordi, J.; Guggiana-Nilo, D.; Bolton, A.D.; Prabha, S.; Ballotti, K.; Herrera, K.; Rennekamp, A.J.; Peterson, R.T.; Lutz, T.A.; Engert, F. High-throughput screening for selective appetite modulators: A multibehavioral and translational drug discovery strategy. *Sci. Adv.* **2018**, *4*, eaav1966, doi:10.1126/sciadv.aav1966.

143. Green, M.H.; Ho, R.K.; Hale, M.E. Movement and function of the pectoral fins of the larval zebrafish (Danio rerio) during slow swimming. *J. Exp. Biol.* **2011**, *214*, 3111–3123, doi:10.1242/jeb.057497.

144. Plaut, I. Effects of fin size on swimming performance, swimming behaviour and routine activity of zebrafish

Danio rerio. *J. Exp. Biol.* **2000**, *203*, 813–820.

145. Lange, M.; Norton, W.; Coolen, M.; Chaminade, M.; Merker, S.; Proft, F.; Schmitt, A.; Vernier, P.; Lesch, K.P.; Bally-Cuif, L. The ADHD-susceptibility gene lphn3.1 modulates dopaminergic neuron formation and locomotor activity during zebrafish development. *Mol. Psychiatry* **2012**, *17*, 946–954, doi:10.1038/mp.2012.29.

146. Huang, J.; Zhong, Z.; Wang, M.; Chen, X.; Tan, Y.; Zhang, S.; He, W.; He, X.; Huang, G.; Lu, H.; et al. Circadian modulation of dopamine levels and dopaminergic neuron development contributes to attention deficiency and hyperactive behavior. *J. Neurosci.* **2015**, *35*, 2572–2587, doi:10.1523/JNEUROSCI.2551-14.2015.

147. Yang, L.; Chang, S.; Lu, Q.; Zhang, Y.; Wu, Z.; Sun, X.; Cao, Q.; Qian, Y.; Jia, T.; Xu, B.; et al. A new locus regulating MICALL2 expression was identified for association with executive inhibition in children with attention deficit hyperactivity disorder. *Mol. Psychiatry* **2018**, *23*, 1014–1020, doi:10.1038/mp.2017.74.

148. Spulber, S.; Kilian, P.; Ibrahim, W.N.W.; Onishchenko, N.; Ulhaq, M.; Norrgren, L.; Negri, S.; Di Tuccio, M.; Ceccatelli, S. PFOS induces behavioral alterations, including spontaneous hyperactivity that is corrected by dexamfetamine in zebrafish larvae. *PLoS One* **2014**, *9*, e94227, doi:10.1371/journal.pone.0094227.

149. Parker, M.O.; Millington, M.E.; Combe, F.J.; Brennan, C.H. Development and implementation of a three-choice serial reaction time task for zebrafish (Danio rerio). *Behav. Brain Res.* **2012**, *227*, 73–80, doi:10.1016/j.bbr.2011.10.037.

150. Parker, M.O.; Brock, A.J.; Sudwarts, A.; Brennan, C.H. Atomoxetine reduces anticipatory responding in a 5-choice serial reaction time task for adult zebrafish. *Psychopharmacology (Berl).* **2014**, *231*, 2671–2679, doi:10.1007/s00213-014-3439-z.

151. Echevarria, D.J.; Jouandot, D.J.; Toms, C.N. Assessing attention in the zebrafish. Are we there yet? *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2011**, *35*, 1416–1420, doi:10.1016/j.pnpbp.2011.01.020.

152. Bushnell, P.J. Behavioral approaches to the assessment of attention in animals. *Psychopharmacology (Berl).* **1998**, *138*, 231–259, doi:10.1007/s002130050668.

153. Abril-De-Abreu, R.; Cruz, J.; Oliveira, R.F. Social Eavesdropping in Zebrafish: Tuning of Attention to Social Interactions. *Sci. Rep.* **2015**, *5*, 12678, doi:10.1038/srep12678.

154. Stednitz, S.J.; McDermott, E.M.; Ncube, D.; Tallafuss, A.; Eisen, J.S.; Washbourne, P. Forebrain Control of Behaviorally Driven Social Orienting in Zebrafish. *Curr. Biol.* **2018**, *28*, 2445-2451.e3, doi:10.1016/j.cub.2018.06.016.

155. Braida, D.; Ponzoni, L.; Martucci, R.; Sala, M. A new model to study visual attention in zebrafish. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2014**, *55*, 80–86, doi:10.1016/j.pnpbp.2014.03.010.

156. Fizet, J.; Cassel, J.C.; Kelche, C.; Meunier, H. A review of the 5-Choice Serial Reaction Time (5-CSRT) task in different vertebrate models. *Neurosci. Biobehav. Rev.* 2016, *71*, 135–153.

157. Parker, M.O.; Gaviria, J.; Haigh, A.; Millington, M.E.; Brown, V.J.; Combe, F.J.; Brennan, C.H. Discrimination reversal and attentional sets in zebrafish (Danio rerio). *Behav. Brain Res.* **2012**, *232*, 264–268, doi:10.1016/j.bbr.2012.04.035.

158. Neri, P. Feature binding in zebrafish. *Anim. Behav.* **2012**, *84*, 485–493, doi:10.1016/j.anbehav.2012.06.005.

159. Arcos-Burgos, M.; Jain, M.; Acosta, M.T.; Shively, S.; Stanescu, H.; Wallis, D.; Domené, S.; Vélez, J.I.; Karkera, J.D.; Balog, J.; et al. A common variant of the latrophilin 3 gene, LPHN3, confers susceptibility to ADHD and predicts effectiveness of stimulant medication. *Mol. Psychiatry* **2010**, *15*, 1053–1066, doi:10.1038/mp.2010.6.

160. Ribasés, M.; Ramos-Quiroga, J.A.; Sánchez-Mora, C.; Bosch, R.; Richarte, V.; Palomar, G.; Gastaminza, X.; Bielsa, A.; Arcos-Burgos, M.; Muenke, M.; et al. Contribution of LPHN3 to the genetic susceptibility to ADHD in adulthood: A replication study. *Genes, Brain Behav.* **2011**, *10*, 149–157, doi:10.1111/j.1601-183X.2010.00649.x.

161. Martinez, A.F.; Abe, Y.; Hong, S.; Molyneux, K.; Yarnell, D.; Löhr, H.; Driever, W.; Acosta, M.T.; Arcos-Burgos, M.; Muenke, M. An Ultraconserved Brain-Specific Enhancer Within ADGRL3 (LPHN3) Underpins Attention-Deficit/Hyperactivity Disorder Susceptibility. *Biol. Psychiatry* **2016**, *80*, 943–954, doi:10.1016/j.biopsych.2016.06.026.

162. Reuter, I.; Knaup, S.; Romanos, M.; Lesch, K.P.; Drepper, C.; Lillesaar, C. Developmental exposure to acetaminophen does not induce hyperactivity in zebrafish larvae. *J. Neural Transm.* **2016**, *123*, 841–848, doi:10.1007/s00702-016-1556-z.

163. Tay, T.L.; Ronneberger, O.; Ryu, S.; Nitschke, R.; Driever, W. Comprehensive catecholaminergic projectome analysis reveals single-neuron integration of zebrafish ascending and descending dopaminergic systems. *Nat. Commun.* **2011**, *2*, 171, doi:10.1038/ncomms1171.

164. Lange, M.; Froc, C.; Grunwald, H.; Norton, W.H.J.; Bally-Cuif, L. Pharmacological analysis of zebrafish lphn3.1 morphant larvae suggests that saturated dopaminergic signaling could underlie the ADHD-like locomotor hyperactivity. *Prog. Neuro-Psychopharmacology Biol. Psychiatry* **2018**, *84*, 181–189, doi:10.1016/j.pnpbp.2018.02.010.

165. Wallis, D.; Hill, D.S.; Mendez, I.A.; Abbott, L.C.; Finnell, R.H.; Wellman, P.J.; Setlow, B. Initial characterization of mice null for Lphn3, a gene implicated in ADHD and addiction. *Brain Res.* **2012**, *1463*, 85–92, doi:10.1016/j.brainres.2012.04.053.

166. Regan, S.L.; Cryan, M.T.; Williams, M.T.; Vorhees, C. V.; Ross, A.E. Enhanced Transient Striatal Dopamine Release and Reuptake in Lphn3 Knockout Rats. *ACS Chem. Neurosci.* **2020**, *11*, 1171–1177, doi:10.1021/acschemneuro.0c00033.

167. Gerlai, R.; Lahav, M.; Guo, S.; Rosenthal, A. Drinks like a fish: Zebra fish (Danio rerio) as a behavior genetic model to study alcohol effects. *Pharmacol. Biochem. Behav.* **2000**, *67*, 773–782, doi:10.1016/S0091-3057(00)00422-6.

168. Wang, T.; Liu, Y.; Liu, H.; Li, C.; Wang, Y. Auriculasin from Flemingia philippinensis roots shows good therapeutic indexes on hyperactive behavior in zebrafish. *Biochem. Biophys. Res. Commun.* **2018**, *503*, 1254–1259, doi:10.1016/j.bbrc.2018.07.033.

169. Unlu, G.; Qi, X.; Gamazon, E.R.; Melville, D.B.; Patel, N.; Rushing, A.R.; Hashem, M.; Al-Faifi, A.; Chen, R.; Li, B.; et al. Phenome-based approach identifies RIC1-linked Mendelian syndrome through zebrafish models, biobank associations and clinical studies. *Nat. Med.* **2020**, *26*, 98–109, doi:10.1038/s41591-019-0705-y.

170. Lasky-Su, J.; Anney, R.J.L.; Neale, B.M.; Franke, B.; Zhou, K.; Maller, J.B.; Vasquez, A.A.; Chen, W.; Asherson, P.; Buitelaar, J.; et al. Genome-wide association scan of the time to onset of attention deficit hyperactivity disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **2008**, *147*, 1355–1358, doi:10.1002/ajmg.b.30869.

171. Lechermeier, C.G.; D'Orazio, A.; Romanos, M.; Lillesaar, C.; Drepper, C. Distribution of transcripts of the GFOD gene family members gfod1 and gfod2 in the zebrafish central nervous system. *Gene Expr. Patterns* **2020**, *36*, 119111, doi:10.1016/j.gep.2020.119111.

172. Laplana, M.; Royo, J.L.; García, L.F.; Aluja, A.; Gomez-Skarmeta, J.L.; Fibla, J. SIRPB1 copy-number polymorphism as candidate quantitative trait locus for impulsive-disinhibited personality. *Genes, Brain Behav.* **2014**, *13*, 653–662, doi:10.1111/gbb.12154.

173. Teles, M.; Oliveira, R.F. Quantifying aggressive behavior in Zebrafish. In *Methods in Molecular Biology*; Humana Press Inc., 2016; Vol. 1451, pp. 293–305.

174. Norton, W.H.; Winberg, S. Studying aggression in zebrafish. In *Behavioral and Neural Genetics of Zebrafish*; Elsevier, 2020; pp. 481–491.

175. Carreño Gutiérrez, H.; Vacca, I.; Schoenmacker, G.; Cleal, M.; Tochwin, A.; O'Connor, B.; Young, A.M.J.; Vasquez, A.A.; Winter, M.J.; Parker, M.O.; et al. Screening for drugs to reduce zebrafish aggression identifies caffeine and sildenafil. *Eur. Neuropsychopharmacol.* **2020**, *30*, 17–29, doi:10.1016/j.euroneuro.2019.10.005.

176. Meshalkina, D.A.; N. Kizlyk, M.; V. Kysil, E.; Collier, A.D.; Echevarria, D.J.; Abreu, M.S.; Barcellos, L.J.G.; Song, C.; Warnick, J.E.; Kyzar, E.J.; et al. Zebrafish models of autism spectrum disorder. *Exp. Neurol.* **2018**, *299*, 207–216, doi:10.1016/j.expneurol.2017.02.004.

177. Norton, W.H.J.; Manceau, L.; Reichmann, F. The visually mediated social preference test: A novel technique to

measure social behavior and behavioral disturbances in Zebrafish. In *Methods in Molecular Biology*; Humana Press Inc., 2019; Vol. 2011, pp. 121–132.

178. Miller, N.Y.; Gerlai, R. Shoaling in zebrafish: What we don't know. *Rev. Neurosci.* **2011**, *22*, 17–25, doi:10.1515/RNS.2011.004.

179. Hoffman, K.L. *What can animal models tell us about depressive disorders?*; 2016; ISBN 9780081000991.

180. Dalla Vecchia, E.; Donato, V. Di; Young, A.M.J.; Bene, F. Del; Norton, W.H.J. Reelin Signaling Controls the Preference for Social Novelty in Zebrafish. *Front. Behav. Neurosci.* **2019**, *13*, 214, doi:10.3389/fnbeh.2019.00214.

181. Liu, C.X.; Li, C.Y.; Hu, C.C.; Wang, Y.; Lin, J.; Jiang, Y.H.; Li, Q.; Xu, X. CRISPR/Cas9-induced shank3b mutant zebrafish display autism-like behaviors. *Mol. Autism* **2018**, *9*, 23, doi:10.1186/s13229-018-0204-x.

182. Webb, K.J.; Norton, W.H.J.; Trümbach, D.; Meijer, A.H.; Ninkovic, J.; Topp, S.; Heck, D.; Marr, C.; Wurst, W.; Theis, F.J.; et al. Zebrafish reward mutants reveal novel transcripts mediating the behavioral effects of amphetamine. *Genome Biol.* **2009**, *10*, R81, doi:10.1186/gb-2009-10-7-r81.

183. García-González, J.; Brock, A.J.; Parker, M.O.; Riley, R.J.; Joliffe, D.; Sudwarts, A.; Teh, M.T.; Busch-Nentwich, E.M.; Stemple, D.L.; Martineau, A.R.; et al. Identification of slit3 as a locus affecting nicotine preference in zebrafish and human smoking behaviour. *Elife* **2020**, *9*, doi:10.7554/eLife.51295.

184. Brenner, R.G.; Oliveri, A.N.; Sinnott-Armstrong, W.; Levin, E.D. Effects of sub-chronic methylphenidate on risk-taking and sociability in zebrafish (Danio rerio). *Naunyn. Schmiedebergs. Arch. Pharmacol.* **2020**, *393*, 1373–1381, doi:10.1007/s00210-020-01835-z.

185. Levin, E.D.; Sledge, D.; Roach, S.; Petro, A.; Donerly, S.; Linney, E. Persistent behavioral impairment caused by embryonic methylphenidate exposure in zebrafish. *Neurotoxicol. Teratol.* **2011**, *33*, 668–673, doi:10.1016/j.ntt.2011.06.004.

186. Donzelli, G.; Carducci, A.; Llopis-Gonzalez, A.; Verani, M.; Llopis-Morales, A.; Cioni, L.; Morales-Suárez-varela, M. The association between lead and attention-deficit/hyperactivity disorder: A systematic review. *Int. J. Environ. Res. Public Health* **2019**, *16*, 382, doi:10.3390/ijerph16030382.

187. Dou, C.; Zhang, J. Effects of lead on neurogenesis during zebrafish embryonic brain development. *J. Hazard. Mater.* **2011**, *194*, 277–282, doi:10.1016/j.jhazmat.2011.07.106.

188. Lovato, A.K.; Creton, R.; Colwill, R.M. Effects of embryonic exposure to polychlorinated biphenyls (PCBs) on larval zebrafish behavior. *Neurotoxicol. Teratol.* **2016**, *53*, 1–10, doi:10.1016/j.ntt.2015.11.002.

189. Liew, Z.; Ritz, B.; Rebordosa, C.; Lee, P.C.; Olsen, J. Acetaminophen use during pregnancy, behavioral problems, and hyperkinetic disorders. *JAMA Pediatr.* **2014**, *168*, 313–320, doi:10.1001/jamapediatrics.2013.4914.

190. Saad, A.; Hegde, S.; Kechichian, T.; Gamble, P.; Rahman, M.; Stutz, S.J.; Anastasio, N.C.; Alshehri, W.; Lei, J.; Mori, S.; et al. Is There a Causal Relation between Maternal Acetaminophen Administration and ADHD? *PLoS One* **2016**, *11*, e0157380, doi:10.1371/journal.pone.0157380.

191. Reuter, M.S.; Riess, A.; Moog, U.; Briggs, T.A.; Chandler, K.E.; Rauch, A.; Stampfer, M.; Steindl, K.; Glaser, D.; Joset, P.; et al. FOXP2 variants in 14 individuals with developmental speech and language disorders broaden the mutational and clinical spectrum. *J Med Genet* **2017**, *54*, 64–72, doi:10.1136/jmedgenet-2016-104094.

192. Faraone, S. V.; Asherson, P.; Banaschewski, T.; Biederman, J.; Buitelaar, J.K.; Ramos-Quiroga, J.A.; Rohde, L.A.; Sonuga-Barke, E.J.S.; Tannock, R.; Franke, B. Attention-deficit/hyperactivity disorder. *Nat. Rev. Dis. Prim.* **2015**, *1*, doi:10.1038/nrdp.2015.20.

193. Bellen, H.J.; Tong, C.; Tsuda, H. 100 years of Drosophila research and its impact on vertebrate neuroscience: A history lesson for the future. *Nat. Rev. Neurosci.* **2010**, *11*, 514–522, doi:10.1038/nrn2839.

194. Benzer, S. BEHAVIORAL MUTANTS OF Drosophila ISOLATED BY COUNTERCURRENT DISTRIBUTION. *Proc. Natl. Acad. Sci.* **1967**, *58*, 1112–1119, doi:10.1073/pnas.58.3.1112.

195. Bier, E. Drosophila, the golden bug, emerges as a tool for human genetics. *Nat. Rev. Genet.* **2005**, *6*, 9–23, doi:10.1038/nrg1503.

196. Burne, T.; Scott, E.; Van Swinderen, B.; Hilliard, M.; Reinhard, J.; Claudianos, C.; Eyles, D.; McGrath, J. Big ideas for small brains: What can psychiatry learn from worms, flies, bees and fish. *Mol. Psychiatry* **2011**, *16*, 7–16, doi:10.1038/mp.2010.35.

197. Coll-Tane, M.; Krebbers, A.; Castells-Nobau, A.; Zweier, C.; Schenck, A. Intellectual disability and autism spectrum disorders "on the fly": Insights from Drosophila. *DMM Dis. Model. Mech.* **2019**, *12*, doi:10.1242/dmm.039180.

198. Hirth, F.; Reichert, H. Conserved genetic programs in insect and mammalian brain development. *Bioessays* **1999**, *21*, 677–684, doi:10.1002/(SICI)1521-1878(199908)21:8<677::AID-BIES7>3.0.CO;2-8.

199. van der Voet, M.; Nijhof, B.; Oortveld, M.A.; Schenck, A. Drosophila models of early onset cognitive disorders and their clinical applications. *Neurosci Biobehav Rev* **2014**, *46 Pt 2*, 326–342, doi:10.1016/j.neubiorev.2014.01.013.

200. Brand, A.H.; Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **1993**, *118*, 401–415.

201. Duffy, J.B. GAL4 system in Drosophila: A fly geneticist's Swiss army knife. *Genesis* **2002**, *34*, 1–15, doi:10.1002/gene.10150.

202. Bellen, H.J.; Levis, R.W.; He, Y.; Carlson, J.W.; Evans-Holm, M.; Bae, E.; Kim, J.; Metaxakis, A.; Savakis, C.; Schulze, K.L.; et al. The Drosophila gene disruption project: Progress using transposons with distinctive site specificities. *Genetics* **2011**, *188*, 731–743, doi:10.1534/genetics.111.126995.

203. Bischof, J.; Björklund, M.; Furger, E.; Schertel, C.; Taipale, J.; Basler, K. A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. *Dev.* **2012**, *140*, 2434–2442, doi:10.1242/dev.088757.

204. Dietzl, G.; Chen, D.; Schnorrer, F.; Su, K.C.; Barinova, Y.; Fellner, M.; Gasser, B.; Kinsey, K.; Oppel, S.; Scheiblauer, S.; et al. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **2007**, *448*, 151–156, doi:10.1038/nature05954.

205. Jenett, A.; Rubin, G.M.; Ngo, T.T.; Shepherd, D.; Murphy, C.; Dionne, H.; Pfeiffer, B.D.; Cavallaro, A.; Hall, D.; Jeter, J.; et al. A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep* **2012**, *2*, 991–1001, doi:10.1016/j.celrep.2012.09.011.

206. Matthews, K.A.; Kaufman, T.C.; Gelbart, W.M. Research resources for Drosophila: the expanding universe. *Nat Rev Genet* **2005**, *6*, 179–193, doi:10.1038/nrg1554.

207. Perkins, L.A.; Holderbaum, L.; Tao, R.; Hu, Y.; Sopko, R.; McCall, K.; Yang-Zhou, D.; Flockhart, I.; Binari, R.; Shim, H.S.; et al. The Transgenic RNAi Project at Harvard Medical School: Resources and Validation. *Genetics* **2015**, *201*, 843–852, doi:10.1534/genetics.115.180208.

208. van Alphen, B.; van Swinderen, B. Drosophila strategies to study psychiatric disorders. *Brain Res Bull* **2013**, *92*, 1–11, doi:10.1016/j.brainresbull.2011.09.007.

209. van der Voet, M.; Harich, B.; Franke, B.; Schenck, A. ADHD-associated dopamine transporter, latrophilin and neurofibromin share a dopamine-related locomotor signature in Drosophila. *Mol Psychiatry* **2016**, *21*, 565–573, doi:10.1038/mp.2015.55.

210. Faville, R.; Kottler, B.; Goodhill, G.J.; Shaw, P.J.; Van Swinderen, B. How deeply does your mutant sleep? Probing arousal to better understand sleep defects in Drosophila. *Sci. Rep.* **2015**, *5*, 8454, doi:10.1038/srep08454.

211. Geissmann, Q.; Garcia Rodriguez, L.; Beckwith, E.J.; French, A.S.; Jamasb, A.R.; Gilestro, G.F. Ethoscopes: An open platform for high-throughput ethomics. *PLoS Biol* **2017**, *15*, e2003026, doi:10.1371/journal.pbio.2003026.

212. Murphy, K.R.; Park, J.H.; Huber, R.; Ja, W.W. Simultaneous measurement of sleep and feeding in individual Drosophila. *Nat Protoc* **2017**, *12*, 2355–2366, doi:10.1038/nprot.2017.096.

213. Branson, K.; Robie, A.A.; Bender, J.; Perona, P.; Dickinson, M.H. High-throughput ethomics in large groups of Drosophila. *Nat. Methods* **2009**, *6*, 451–457, doi:10.1038/nmeth.1328.

214. Kabra, M.; Robie, A.A.; Rivera-Alba, M.; Branson, S.; Branson, K. JAABA: interactive machine learning for automatic annotation of animal behavior. *Nat Methods* **2013**, *10*, 64–67, doi:10.1038/nmeth.2281.

215. Frighetto, G.; Zordan, M.A.; Castiello, U.; Megighian, A. Action-based attention in drosophila melanogaster. *J. Neurophysiol.* **2019**, *121*, 2428–2432, doi:10.1152/jn.00164.2019.

216. Qiao, B.; Li, C.; Allen, V.W.; Shirasu-Hiza, M.; Syed, S. Automated analysis of long-term grooming behavior in Drosophila using a k-nearest neighbors classifier. *Elife* **2018**, *7*, doi:10.7554/eLife.34497.

217. Luck, S.J.; Gold, J.M. The construct of attention in schizophrenia. *Biol Psychiatry* **2008**, *64*, 34–39, doi:10.1016/j.biopsych.2008.02.014.

218. McClain, M.B.; Hasty Mills, A.M.; Murphy, L.E. Inattention and hyperactivity/impulsivity among children with attention-deficit/hyperactivity-disorder, autism spectrum disorder, and intellectual disability. *Res Dev Disabil* **2017**, *70*, 175–184, doi:10.1016/j.ridd.2017.09.009.

219. De Bivort, B.L.; Van Swinderen, B. Evidence for selective attention in the insect brain. *Curr. Opin. Insect Sci.* **2016**, *15*, 9–15, doi:10.1016/j.cois.2016.02.007.

220. van Swinderen, B. Attention in Drosophila. *Int Rev Neurobiol* **2011**, *99*, 51–85, doi:10.1016/B978-0-12-387003-2.00003-3.

221. Gotz, K.G. Visual guidance in Drosophila. *Basic Life Sci* **1980**, *16*, 391–407, doi:10.1007/978-1-4684-7968-3_28.

222. Paulk, A.C.; Kirszenblat, L.; Zhou, Y.; van Swinderen, B. Closed-Loop Behavioral Control Increases Coherence in the Fly Brain. *J. Neurosci.* **2015**, *35*, 10304–10315, doi:10.1523/Jneurosci.0691-15.2015.

223. Koenig, S.; Wolf, R.; Heisenberg, M. Visual Attention in Flies-Dopamine in the Mushroom Bodies Mediates the After-Effect of Cueing. *PLoS One* **2016**, *11*, e0161412, doi:10.1371/journal.pone.0161412.

224. Van Swinderen, B.; Andretic, R. Dopamine in Drosophila: setting arousal thresholds in a miniature brain. *Proc Biol Sci* **2011**, *278*, 906–913, doi:10.1098/rspb.2010.2564.

225. Colomb, J.; Reiter, L.; Blaszkiewicz, J.; Wessnitzer, J.; Brembs, B. Open source tracking and analysis of adult Drosophila locomotion in Buridan's paradigm with and without visual targets. *PLoS One* **2012**, *7*, e42247, doi:10.1371/journal.pone.0042247.

226. Ferguson, L.; Petty, A.; Rohrscheib, C.; Troup, M.; Kirszenblat, L.; Eyles, D.W.; van Swinderen, B. Transient dysregulation of dopamine signaling in a developing Drosophila arousal circuit permanently impairs behavioral responsiveness in adults. *Front. Psychiatry* **2017**, *8*, 22, doi:10.3389/fpsyt.2017.00022.

227. Kirszenblat, L.; Ertekin, D.; Goodsell, J.; Zhou, Y.; Shaw, P.J.; van Swinderen, B. Sleep regulates visual selective attention in Drosophila. *J Exp Biol* **2018**, *221*, doi:10.1242/jeb.191429.

228. Seelig, J.D.; Chiappe, M.E.; Lott, G.K.; Dutta, A.; Osborne, J.E.; Reiser, M.B.; Jayaraman, V. Two-photon calcium imaging from head-fixed Drosophila during optomotor walking behavior. *Nat Methods* **2010**, *7*, 535–540, doi:10.1038/nmeth.1468.

229. Tang, S.; Juusola, M. Intrinsic activity in the fly brain gates visual information during behavioral choices. *PLoS One* **2010**, *5*, e14455, doi:10.1371/journal.pone.0014455.

230. van Swinderen, B.; Brembs, B. Attention-like deficit and hyperactivity in a Drosophila memory mutant. *J Neurosci* **2010**, *30*, 1003–1014, doi:10.1523/JNEUROSCI.4516-09.2010.

231. Koenig, S.; Wolf, R.; Heisenberg, M. Vision in Flies: Measuring the Attention Span. *PLoS One* **2016**, *11*, e0148208, doi:10.1371/journal.pone.0148208.

232. Dalley, J.W.; Everitt, B.J.; Robbins, T.W. Impulsivity, Compulsivity, and Top-Down Cognitive Control. *Neuron* **2011**, *69*, 680–694, doi:10.1016/j.neuron.2011.01.020.

233. Lee, H.G.; Kim, Y.C.; Dunning, J.S.; Han, K.A. Recurring ethanol exposure induces disinhibited courtship in Drosophila. *PLoS One* **2008**, *3*, e1391, doi:10.1371/journal.pone.0001391.

234. Aranda, G.P.; Hinojos, S.J.; Sabandal, P.R.; Evans, P.D.; Han, K.A. Behavioral sensitization to the disinhibition effect of ethanol requires the dopamine/ecdysone receptor in drosophila. *Front. Syst. Neurosci.* **2017**, *11*, 56, doi:10.3389/fnsys.2017.00056.

235. Cook, E.H.; Stein, M.A.; Krasowski, M.D.; Cox, N.J.; Olkon, D.M.; Kieffer, J.E.; Leventhal, B.L. Association of attention-deficit disorder and the dopamine transporter gene. *Am. J. Hum. Genet.* **1995**, *56*, 993–998.

236. van Dyck, C.H.; Malison, R.T.; Jacobsen, L.K.; Seibyl, J.P.; Staley, J.K.; Laruelle, M.; Baldwin, R.M.; Innis, R.B.; Gelernter, J. Increased dopamine transporter availability associated with the 9-repeat allele of the SLC6A3 gene. *J Nucl Med* **2005**, *46*, 745–751.

237. Hamilton, P.J.; Campbell, N.G.; Sharma, S.; Erreger, K.; Herborg Hansen, F.; Saunders, C.; Belovich, A.N.; Consortium, N.A.A.S.; Sahai, M.A.; Cook, E.H.; et al. De novo mutation in the dopamine transporter gene associates dopamine dysfunction with autism spectrum disorder. *Mol Psychiatry* **2013**, *18*, 1315–1323, doi:10.1038/mp.2013.102.

238. Kume, K.; Kume, S.; Park, S.K.; Hirsh, J.; Jackson, F.R. Dopamine is a regulator of arousal in the fruit fly. *J Neurosci* **2005**, *25*, 7377–7384, doi:10.1523/JNEUROSCI.2048-05.2005.

239. Landgraf, D.; Joiner, W.J.; McCarthy, M.J.; Kiessling, S.; Barandas, R.; Young, J.W.; Cermakian, N.; Welsh, D.K. The mood stabilizer valproic acid opposes the effects of dopamine on circadian rhythms. *Neuropharmacology* **2016**, *107*, 262–270, doi:10.1016/j.neuropharm.2016.03.047.

240. Kayser, M.S.; Yue, Z.; Sehgal, A. A critical period of sleep for development of courtship circuitry and behavior in Drosophila. *Science (80-. ).* **2014**, *344*, 269–274, doi:10.1126/science.1250553.

241. King, L.B.; Koch, M.; Murphy, K.R.; Velazquez, Y.; Ja, W.W.; Tomchik, S.M. Neurofibromin Loss of Function Drives Excessive Grooming in Drosophila. *G3* **2016**, *6*, 1083–1093, doi:10.1534/g3.115.026484.

242. Barker, D.; Wright, E.; Nguyen, K.; Cannon, L.; Fain, P.; Goldgar, D.; Bishop, D.T.; Carey, J.; Baty, B.; Kivlin, J.; et al. Gene for von Recklinghausen neurofibromatosis is in the pericentromeric region of chromosome 17. *Science (80-. ).* **1987**, *236*, 1100–1102, doi:10.1126/science.3107130.

243. Mautner, V.F.; Kluwe, L.; Thakker, S.D.; Leark, R.A. Treatment of ADHD in neurofibromatosis type 1. *Dev Med Child Neurol* **2002**, *44*, 164–170, doi:10.1017/s0012162201001876.

244. Payne, J.M.; Haebich, K.M.; MacKenzie, R.; Walsh, K.S.; Hearps, S.J.C.; Coghill, D.; Barton, B.; Pride, N.A.; Ullrich, N.J.; Tonsgard, J.H.; et al. Cognition, ADHD Symptoms, and Functional Impairment in Children and Adolescents With Neurofibromatosis Type 1. *J Atten Disord* **2019**, 1087054719894384, doi:10.1177/1087054719894384.

245. Pride, N.A.; Payne, J.M.; North, K.N. The impact of ADHD on the cognitive and academic functioning of children with NF1. *Dev Neuropsychol* **2012**, *37*, 590–600, doi:10.1080/87565641.2012.695831.

246. Orsini, C.A.; Setlow, B.; Dejesus, M.; Galaviz, S.; Loesch, K.; Ioerger, T.; Wallis, D. Behavioral and transcriptomic profiling of mice null for Lphn3, a gene implicated in ADHD and addiction. *Mol. Genet. Genomic Med.* **2016**, *4*, 322–343, doi:10.1002/mgg3.207.

247. Vilidaite, G.; Norcia, A.M.; West, R.J.H.; Elliott, C.J.H.; Pei, F.; Wade, A.R.; Baker, D.H. Autism sensory dysfunction in an evolutionarily conserved system. *Proc Biol Sci* **2018**, *285*, 20182255, doi:10.1098/rspb.2018.2255.

248. Rohde, P.D.; Madsen, L.S.; Neumann Arvidson, S.M.; Loeschcke, V.; Demontis, D.; Kristensen, T.N. Testing candidate genes for attention-deficit/hyperactivity disorder in fruit flies using a high throughput assay for complex behavior. *Fly* **2016**, *10*, 25–34, doi:10.1080/19336934.2016.1158365.

249. Lai, C.S.; Fisher, S.E.; Hurst, J.A.; Vargha-Khadem, F.; Monaco, A.P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **2001**, *413*, 519–523, doi:10.1038/35097076.

250. Castells-Nobau, A.; Eidhof, I.; Fenckova, M.; Brenman-Suttner, D.B.; Scheffer-De Gooyert, J.M.; Christine, S.; Schellevis, R.L.; Van Der Laan, K.; Quentin, C.; Van Ninhuijs, L.; et al. Conserved regulation of neurodevelopmental processes and behavior by FoxP in Drosophila. *PLoS One* **2019**, *14*, e0211652, doi:10.1371/journal.pone.0211652.

251. DasGupta, S.; Ferreira, C.H.; Miesenböck, G. FoxP influences the speed and accuracy of a perceptual decision in Drosophila. *Science (80-. ).* **2014**, *344*, 901–904, doi:10.1126/science.1252114.

252. Lawton, K.J.; Wassmer, T.L.; Deitcher, D.L. Conserved role of Drosophila melanogaster FoxP in motor coordination and courtship song. *Behav Brain Res* **2014**, *268*, 213–221, doi:10.1016/j.bbr.2014.04.009.

253. Mendoza, E.; Colomb, J.; Rybak, J.; Pfluger, H.J.; Zars, T.; Scharff, C.; Brembs, B. Drosophila FoxP mutants are deficient in operant self-learning. *PLoS One* **2014**, *9*, e100648, doi:10.1371/journal.pone.0100648.

254. Chen, Y.C.; Kuo, H.Y.; Bornschein, U.; Takahashi, H.; Chen, S.Y.; Lu, K.M.; Yang, H.Y.; Chen, G.M.; Lin, J.R.; Lee, Y.H.; et al. Foxp2 controls synaptic wiring of corticostriatal circuits and vocal communication by opposing Mef2c. *Nat. Neurosci.* **2016**, *19*, 1513–1522, doi:10.1038/nn.4380.

255. Blanchard, F.J.; Collins, B.; Cyran, S.A.; Hancock, D.H.; Taylor, M. V.; Blau, J. The transcription factor Mef2 is required for normal circadian behavior in Drosophila. *J. Neurosci.* **2010**, *30*, 5855–5865, doi:10.1523/JNEUROSCI.2688-09.2010.

256. Sivachenko, A.; Li, Y.; Abruzzi, K.C.; Rosbash, M. The transcription factor Mef2 links the Drosophila core clock to Fas2, neuronal morphology, and circadian behavior. *Neuron* **2013**, *79*, 281–292, doi:10.1016/j.neuron.2013.05.015.

257. Klein, M.; Singgih, E.L.; van Rens, A.; Demontis, D.; Børglum, A.D.; Mota, N.R.; Castells-Nobau, A.; Kiemeney, L.A.; Brunner, H.G.; Arias-Vasquez, A.; et al. Contribution of intellectual disability–related genes to ADHD risk and to locomotor activity in Drosophila. *Am. J. Psychiatry* **2020**, *177*, doi:10.1176/appi.ajp.2019.18050599.

258. Harich, B.; Klein, M.; Ockeloen, C.W.; van der Voet, M.; Schimmel-Naber, M.; de Leeuw, N.; Schenck, A.; Franke, B. From man to fly - convergent evidence links FBXO25 to ADHD and comorbid psychiatric phenotypes. *J Child Psychol Psychiatry* **2020**, *61*, 545–555, doi:10.1111/jcpp.13161.

259. Rankin, C.H.; Abrams, T.; Barry, R.J.; Bhatnagar, S.; Clayton, D.F.; Colombo, J.; Coppola, G.; Geyer, M.A.; Glanzman, D.L.; Marsland, S.; et al. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol Learn Mem* **2009**, *92*, 135–138, doi:10.1016/j.nlm.2008.09.012.

260. Biederman, J. Attention-deficit/hyperactivity disorder: A selective overview. *Biol. Psychiatry* **2005**, *57*, 1215–1220, doi:10.1016/j.biopsych.2004.10.020.

261. Faraone, S. V.; Biederman, J.; Spencer, T.; Wilens, T.; Seidman, L.J.; Mick, E.; Doyle, A.E. Attention-deficit/hyperactivity disorder in adults: An overview. *Biol. Psychiatry* **2000**, *48*, 9–20, doi:10.1016/S0006-3223(00)00889-1.

262. Jansiewicz, E.M.; Newschaffer, C.J.; Denckla, M.B.; Mostofsky, S.H. Impaired habituation in children with attention deficit hyperactivity disorder. *Cogn Behav Neurol* **2004**, *17*, 1–8, doi:10.1097/00146965-200403000-00001.

263. Massa, J.; O'Desky, I.H. Impaired visual habituation in adults with ADHD. *J Atten Disord* **2012**, *16*, 553–561, doi:10.1177/1087054711423621.

264. Ramaswami, M. Network plasticity in adaptive filtering and behavioral habituation. *Neuron* **2014**, *82*, 1216–1229, doi:10.1016/j.neuron.2014.04.035.

265. Lisman, J. Excitation, inhibition, local oscillations, or large-scale loops: what causes the symptoms of schizophrenia? *Curr Opin Neurobiol* **2012**, *22*, 537–544, doi:10.1016/j.conb.2011.10.018.

266. Rubenstein, J.L.R.; Merzenich, M.M. Model of autism: Increased ratio of excitation/inhibition in key neural systems. *Genes, Brain Behav.* **2003**, *2*, 255–267, doi:10.1034/j.1601-183X.2003.00037.x.

267. Acevedo, S.F.; Froudarakis, E.I.; Kanellopoulos, A.; Skoulakis, E.M.C. Protection from premature habituation

requires functional mushroom bodies in Drosophila. *Learn. Mem.* **2007**, *14*, 376–384, doi:10.1101/lm.566007.

268. Asztalos, Z.; Arora, N.; Tully, T. Olfactory jump reflex habituation in Drosophila and effects of classical conditioning mutations. *J. Neurogenet.* **2007**, *21*, 1–18, doi:10.1080/01677060701247508.

269. Das, S.; Sadanandappa, M.K.; Dervan, A.; Larkin, A.; Lee, J.A.; Sudhakaran, I.P.; Priya, R.; Heidari, R.; Holohan, E.E.; Pimentel, A.; et al. Plasticity of local GABAergic interneurons drives olfactory habituation. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, E646-54, doi:10.1073/pnas.1106411108.

270. Paranjpe, P.; Rodrigues, V.; VijayRaghavan, K.; Ramaswami, M. Gustatory habituation in Drosophila relies on rutabaga (adenylate cyclase)-dependent plasticity of GABAergic inhibitory neurons. *Learn Mem* **2012**, *19*, 627–635, doi:10.1101/lm.026641.112.

271. Engel, J.E.; Wu, C.F. Altered habituation of an identified escape circuit in Drosophila memory mutants. *J. Neurosci.* **1996**, *16*, 3486–3499, doi:10.1523/jneurosci.16-10-03486.1996.

272. Fenckova, M.; Blok, L.E.R.; Asztalos, L.; Goodman, D.P.; Cizek, P.; Singgih, E.L.; Glennon, J.C.; IntHout, J.; Zweier, C.; Eichler, E.E.; et al. Habituation Learning Is a Widely Affected Mechanism in Drosophila Models of Intellectual Disability and Autism Spectrum Disorders. *Biol. Psychiatry* **2019**, *86*, 294–305, doi:10.1016/j.biopsych.2019.04.029.

273. Mullin, A.P.; Sadanandappa, M.K.; Ma, W.; Dickman, D.K.; VijayRaghavan, K.; Ramaswami, M.; Sanyal, S.; Faundez, V. Gene dosage in the dysbindin schizophrenia susceptibility network differentially affect synaptic function and plasticity. *J Neurosci* **2015**, *35*, 325–338, doi:10.1523/JNEUROSCI.3542-14.2015.

274. Stessman, H.A.; Xiong, B.; Coe, B.P.; Wang, T.; Hoekzema, K.; Fenckova, M.; Kvarnung, M.; Gerdts, J.; Trinh, S.; Cosemans, N.; et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat Genet* **2017**, *49*, 515–526, doi:10.1038/ng.3792.

275. Wolf, F.W.; Eddison, M.; Lee, S.; Cho, W.; Heberlein, U. GSK-3/Shaggy regulates olfactory habituation in Drosophila. *Proc Natl Acad Sci U S A* **2007**, *104*, 4653–4657, doi:10.1073/pnas.0700493104.

276. Schwarz, E.; Tost, H.; Meyer-Lindenberg, A. Working memory genetics in schizophrenia and related disorders: An RDoC perspective. *Am J Med Genet B Neuropsychiatr Genet* **2016**, *171B*, 121–131, doi:10.1002/ajmg.b.32353.

277. Matt Alderson, R.; Kasper, L.J.; Hudec, K.L.; Patros, C.H.G. Attention-deficit/hyperactivity disorder (ADHD) and working memory in adults: A meta-analytic review. *Neuropsychology* 2013, *27*, 287–302.

278. Martinussen, R.; Hayden, J.; Hogg-Johnson, S.; Tannock, R. A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **2005**, *44*, 377–384, doi:10.1097/01.chi.0000153228.72591.73.

279. Neuser, K.; Triphan, T.; Mronz, M.; Poeck, B.; Strauss, R. Analysis of a spatial orientation memory in Drosophila. *Nature* **2008**, *453*, 1244–1247, doi:10.1038/nature07003.

280. Kuntz, S.; Poeck, B.; Sokolowski, M.B.; Strauss, R. The visual orientation memory of Drosophila requires Foraging (PKG) upstream of Ignorant (RSK2) in ring neurons of the central complex. *Learn Mem* **2012**, *19*, 337–340, doi:10.1101/lm.026369.112.

281. Freudenberg, F.; Alttoa, A.; Reif, A. Neuronal nitric oxide synthase (NOS1) and its adaptor, NOS1AP, as a genetic risk factors for psychiatric disorders. *Genes, Brain Behav.* **2015**, *14*, 46–63, doi:10.1111/gbb.12193.

282. Kuntz, S.; Poeck, B.; Strauss, R. Visual Working Memory Requires Permissive and Instructive NO/cGMP Signaling at Presynapses in the Drosophila Central Brain. *Curr Biol* **2017**, *27*, 613–623, doi:10.1016/j.cub.2016.12.056.

283. Cleal, M.; Fontana, B.D.; Ranson, D.C.; McBride, S.D.; Swinny, J.D.; Redhead, E.S.; Parker, M.O. The Free-movement pattern Y-maze: A cross-species measure of working memory and executive function. *Behav. Res. Methods* **2020**, doi:10.3758/s13428-020-01452-x.

284. Corkum, P.; Tannock, R.; Moldofsky, H. Sleep disturbances in children with attention-deficit/hyperactivity

disorder. *J. Am. Acad. Child Adolesc. Psychiatry* **1998**, *37*, 637–646, doi:10.1097/00004583-199806000-00014.

285.	Hodgkins, P.; Setyawan, J.; Mitra, D.; Davis, K.; Quintero, J.; Fridman, M.; Shaw, M.; Harpin, V. Management of ADHD in children across Europe: patient demographics, physician characteristics and treatment patterns. *Eur J Pediatr* **2013**, *172*, 895–906, doi:10.1007/s00431-013-1969-8.

286.	Robinson-Shelton, A.; Malow, B.A. Sleep Disturbances in Neurodevelopmental Disorders. *Curr Psychiatry Rep* **2016**, *18*, 6, doi:10.1007/s11920-015-0638-1.

287.	Ranum, B.M.; Wichstrom, L.; Pallesen, S.; Falch-Madsen, J.; Halse, M.; Steinsbekk, S. Association Between Objectively Measured Sleep Duration and Symptoms of Psychiatric Disorders in Middle Childhood. *JAMA Netw Open* **2019**, *2*, e1918281, doi:10.1001/jamanetworkopen.2019.18281.

288.	Phillips, N.L.; Moore, T.; Teng, A.; Brookes, N.; Palermo, T.M.; Lah, S. Behavioral Interventions for Sleep Disturbances in Children with Neurological and Neurodevelopmental Disorders: A Systematic Review and Meta-analysis of Randomized Controlled Trials. *Sleep* **2020**, doi:10.1093/sleep/zsaa040.

289.	Hendricks, J.C.; Finn, S.M.; Panckeri, K.A.; Chavkin, J.; Williams, J.A.; Sehgal, A.; Pack, A.I. Rest in Drosophila is a sleep-like state. *Neuron* **2000**, *25*, 129–138, doi:10.1016/s0896-6273(00)80877-6.

290.	Shaw, P.J.; Cirelli, C.; Greenspan, R.J.; Tononi, G. Correlates of sleep and waking in Drosophila melanogaster. *Science (80-. ).* **2000**, *287*, 1834–1837, doi:10.1126/science.287.5459.1834.

291.	Ly, S.; Pack, A.I.; Naidoo, N. The neurobiological basis of sleep: Insights from Drosophila. *Neurosci Biobehav Rev* **2018**, *87*, 67–86, doi:10.1016/j.neubiorev.2018.01.015.

292.	Huber, R.; Hill, S.L.; Holladay, C.; Biesiadecki, M.; Tononi, G.; Cirelli, C. Sleep homeostasis in Drosophila melanogaster. *Sleep* **2004**, *27*, 628–639, doi:10.1093/sleep/27.4.628.

293.	Larkin, A.; Chen, M.Y.; Kirszenblat, L.; Reinhard, J.; van Swinderen, B.; Claudianos, C. Neurexin-1 regulates sleep and synaptic plasticity in Drosophila melanogaster. *Eur J Neurosci* **2015**, *42*, 2455–2466, doi:10.1111/ejn.13023.

294.	Tong, H.; Li, Q.; Zhang, Z.C.; Li, Y.; Han, J. Neurexin regulates nighttime sleep by modulating synaptic transmission. *Sci Rep* **2016**, *6*, 38246, doi:10.1038/srep38246.

295.	Sawamura, N.; Ando, T.; Maruyama, Y.; Fujimuro, M.; Mochizuki, H.; Honjo, K.; Shimoda, M.; Toda, H.; Sawamura-Yamamoto, T.; Makuch, L.A.; et al. Nuclear DISC1 regulates CRE-mediated gene transcription and sleep homeostasis in the fruit fly. *Mol Psychiatry* **2008**, *13*, 1069,1138-1148, doi:10.1038/mp.2008.101.

296.	Jagannath, A.; Peirson, S.N.; Foster, R.G. Sleep and circadian rhythm disruption in neuropsychiatric illness. *Curr Opin Neurobiol* **2013**, *23*, 888–894, doi:10.1016/j.conb.2013.03.008.

297.	Menet, J.S.; Rosbash, M. When brain clocks lose track of time: cause or consequence of neuropsychiatric disorders. *Curr Opin Neurobiol* **2011**, *21*, 849–857, doi:10.1016/j.conb.2011.06.008.

298.	Top, D.; Young, M.W. Coordination between differentially regulated circadian clocks generates rhythmic behavior. *Cold Spring Harb. Perspect. Biol.* **2018**, *10*, a033589, doi:10.1101/cshperspect.a033589.

299.	Bargiello, T.A.; Young, M.W. Molecular genetics of a biological clock in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 2142–2146, doi:10.1073/pnas.81.7.2142.

300.	Dubowy, C.; Sehgal, A. Circadian rhythms and sleep in Drosophila melanogaster. *Genetics* **2017**, *205*, 1373–1397, doi:10.1534/genetics.115.185157.

301.	Harker, J.E. The Effect of a Biological Clock on the Developmental Rate of Drosophila Pupae. *J Exp Biol* **1965**, *42*, 323–337.

302.	Konopka, R.J.; Benzer, S. Clock mutants of Drosophila melanogaster. *Proc Natl Acad Sci U S A* **1971**, *68*, 2112–2116, doi:10.1073/pnas.68.9.2112.

303.	Dubruille, R.; Emery, P. A plastic clock: How circadian rhythms respond to environmental cues in drosophila. *Mol. Neurobiol.* **2008**, *38*, 129–145, doi:10.1007/s12035-008-8035-y.

304. Lowenstein, E.G.; Velazquez-Ulloa, N.A. A Fly's Eye View of Natural and Drug Reward. *Front Physiol* **2018**, *9*, 407, doi:10.3389/fphys.2018.00407.

305. Bubeníková-Valešová, V.; Horáček, J.; Vrajová, M.; Höschl, C. Models of schizophrenia in humans and animals based on inhibition of NMDA receptors. *Neurosci. Biobehav. Rev.* **2008**, *32*, 1014–1023, doi:10.1016/j.neubiorev.2008.03.012.

306. Whitehouse, C.M.; Lewis, M.H. Repetitive Behavior in Neurodevelopmental Disorders: Clinical and Translational Findings. *Behav Anal* **2015**, *38*, 163–178, doi:10.1007/s40614-015-0029-2.

307. Seeds, A.M.; Ravbar, P.; Chung, P.; Hampel, S.; Midgley Jr., F.M.; Mensh, B.D.; Simpson, J.H. A suppression hierarchy among competing motor programs drives sequential grooming in Drosophila. *Elife* **2014**, *3*, e02951, doi:10.7554/eLife.02951.

308. Barradale, F.; Sinha, K.; Lebestky, T. Quantification of drosophila grooming behavior. *J. Vis. Exp.* **2017**, *2017*, doi:10.3791/55231.

309. Phillis, R.W.; Bramlage, A.T.; Wotus, C.; Whittaker, A.; Gramates, L.S.; Seppala, D.; Farahanchi, F.; Caruccio, P.; Murphey, R.K. Isolation of mutations affecting neural circuitry required for grooming behavior in Drosophila melanogaster. *Genetics* **1993**, *133*, 581–592.

310. Pitmon, E.; Stephens, G.; Parkhurst, S.J.; Wolf, F.W.; Kehne, G.; Taylor, M.; Lebestky, T. The D1 family dopamine receptor, DopR, potentiates hind leg grooming behavior in Drosophila. *Genes Brain Behav* **2016**, *15*, 327–334, doi:10.1111/gbb.12264.

311. Tauber, J.M.; Vanlandingham, P.A.; Zhang, B. Elevated levels of the vesicular monoamine transporter and a novel repetitive behavior in the Drosophila model of fragile X syndrome. *PLoS One* **2011**, *6*, e27100, doi:10.1371/journal.pone.0027100.

312. Hall, J.; Trent, S.; Thomas, K.L.; O'Donovan, M.C.; Owen, M.J. Genetic risk for schizophrenia: convergence on synaptic pathways involved in plasticity. *Biol Psychiatry* **2015**, *77*, 52–58, doi:10.1016/j.biopsych.2014.07.011.

313. van Bokhoven, H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet* **2011**, *45*, 81–104, doi:10.1146/annurev-genet-110410-132512.

314. Zoghbi, H.Y.; Bear, M.F. Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harb Perspect Biol* **2012**, *4*, doi:10.1101/cshperspect.a009886.

315. Van Berlekom, A.B.; Muflihah, C.H.; Snijders, G.J.L.J.; MacGillavry, H.D.; Middeldorp, J.; Hol, E.M.; Kahn, R.S.; De Witte, L.D. Synapse Pathology in Schizophrenia: A Meta-analysis of Postsynaptic Elements in Postmortem Brain Studies. *Schizophr. Bull.* **2020**, *46*, 374–386, doi:10.1093/schbul/sbz060.

316. Fiala, J.C.; Spacek, J.; Harris, K.M. Dendritic spine pathology: Cause or consequence of neurological disorders? *Brain Res. Rev.* **2002**, *39*, 29–54, doi:10.1016/S0165-0173(02)00158-3.

317. Varghese, M.; Keshav, N.; Jacot-Descombes, S.; Warda, T.; Wicinski, B.; Dickstein, D.L.; Harony-Nicolas, H.; De Rubeis, S.; Drapeau, E.; Buxbaum, J.D.; et al. Autism spectrum disorder: neuropathology and animal models. *Acta Neuropathol* **2017**, *134*, 537–566, doi:10.1007/s00401-017-1736-4.

318. Giros, B.; Caron, M.G. Molecular characterization of the dopamine transporter. *Trends Pharmacol Sci* **1993**, *14*, 43–49, doi:10.1016/0165-6147(93)90029-j.

319. O'Sullivan, M.L.; Martini, F.; von Daake, S.; Comoletti, D.; Ghosh, A. LPHN3, a presynaptic adhesion-GPCR implicated in ADHD, regulates the strength of neocortical layer 2/3 synaptic input to layer 5. *Neural Dev* **2014**, *9*, 7, doi:10.1186/1749-8104-9-7.

320. Rivero, O.; Selten, M.M.; Sich, S.; Popp, S.; Bacmeister, L.; Amendola, E.; Negwer, M.; Schubert, D.; Proft, F.; Kiser, D.; et al. Cadherin-13, a risk gene for ADHD and comorbid disorders, impacts GABAergic function in hippocampus and cognition. *Transl Psychiatry* **2015**, *5*, e655, doi:10.1038/tp.2015.152.

321. Sollner, T.; Whiteheart, S.W.; Brunner, M.; Erdjument-Bromage, H.; Geromanos, S.; Tempst, P.; Rothman, J.E. SNAP receptors implicated in vesicle targeting and fusion. *Nature* **1993**, *362*, 318–324, doi:10.1038/362318a0.

322. Mooney, M.A.; McWeeney, S.K.; Faraone, S. V; Hinney, A.; Hebebrand, J.; Consortium, I.; German, A.G.G.; Nigg, J.T.; Wilmot, B. Pathway analysis in attention deficit hyperactivity disorder: An ensemble approach. *Am J Med Genet B Neuropsychiatr Genet* **2016**, *171*, 815–826, doi:10.1002/ajmg.b.32446.

323. Thapar, A.; Martin, J.; Mick, E.; Arias Vasquez, A.; Langley, K.; Scherer, S.W.; Schachar, R.; Crosbie, J.; Williams, N.; Franke, B.; et al. Psychiatric gene discoveries shape evidence on ADHD's biology. *Mol Psychiatry* **2016**, *21*, 1202–1207, doi:10.1038/mp.2015.163.

324. Collins, C.A.; DiAntonio, A. Synaptic development: insights from Drosophila. *Curr. Opin. Neurobiol.* **2007**, *17*, 35–42, doi:10.1016/j.conb.2007.01.001.

325. Johansen, J.; Halpern, M.E.; Johansen, K.M.; Keshishian, H. Stereotypic morphology of glutamatergic synapses on identified muscle cells of Drosophila larvae. *J Neurosci* **1989**, *9*, 710–725.

326. Furukubo-Tokunaga, K.; Kurita, K.; Honjo, K.; Pandey, H.; Ando, T.; Takayama, K.; Arai, Y.; Mochizuki, H.; Ando, M.; Kamiya, A.; et al. DISC1 causes associative memory and neurodevelopmental defects in fruit flies. *Mol. Psychiatry* **2016**, *21*, 1232–1243, doi:10.1038/mp.2016.15.

327. Pandey, H.; Bourahmoune, K.; Honda, T.; Honjo, K.; Kurita, K.; Sato, T.; Sawa, A.; Furukubo-Tokunaga, K. Genetic interaction of DISC1 and Neurexin in the development of fruit fly glutamatergic synapses. *NPJ Schizophr* **2017**, *3*, 39, doi:10.1038/s41537-017-0040-6.

328. Kaufmann, W.E.; Moser, H.W. Dendritic anomalies in disorders associated with mental retardation. *Cereb Cortex* **2000**, *10*, 981–991, doi:10.1093/cercor/10.10.981.

329. Kulkarni, V.A.; Firestein, B.L. The dendritic tree and brain disorders. *Mol Cell Neurosci* **2012**, *50*, 10–20, doi:10.1016/j.mcn.2012.03.005.

330. Poelmans, G.; Pauls, D.L.; Buitelaar, J.K.; Franke, B. Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. *Am J Psychiatry* **2011**, *168*, 365–377, doi:10.1176/appi.ajp.2010.10070948.

331. Grueber, W.B.; Jan, L.Y.; Jan, Y.N. Tiling of the Drosophila epidermis by multidendritic sensory neurons. *Development* **2002**, *129*, 2867–2878.

332. Jan, Y.N.; Jan, L.Y. Branching out: mechanisms of dendritic arborization. *Nat Rev Neurosci* **2010**, *11*, 316–328, doi:10.1038/nrn2836.

333. O'Brien, S.; Ng-Cordell, E.; Study, D.D.D.; Astle, D.E.; Scerif, G.; Baker, K. STXBP1-associated neurodevelopmental disorder: a comparative study of behavioural characteristics. *J Neurodev Disord* **2019**, *11*, 17, doi:10.1186/s11689-019-9278-9.

334. Stamberger, H.; Nikanorova, M.; Willemsen, M.H.; Accorsi, P.; Angriman, M.; Baier, H.; Benkel-Herrenbrueck, I.; Benoit, V.; Budetta, M.; Caliebe, A.; et al. STXBP1 encephalopathy: A neurodevelopmental disorder including epilepsy. *Neurology* **2016**, *86*, 954–962, doi:10.1212/WNL.0000000000002457.

335. Peng, Y.; Lee, J.; Rowland, K.; Wen, Y.; Hua, H.; Carlson, N.; Lavania, S.; Parrish, J.Z.; Kim, M.D. Regulation of dendrite growth and maintenance by exocytosis. *J Cell Sci* **2015**, *128*, 4279–4292, doi:10.1242/jcs.174771.

336. Gladstone, M.; Su, T.T. Chemical genetics and drug screening in Drosophila cancer models. *J Genet Genomics* **2011**, *38*, 497–504, doi:10.1016/j.jgg.2011.09.003.

337. Pandey, U.B.; Nichols, C.D. Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacol Rev* **2011**, *63*, 411–436, doi:10.1124/pr.110.003293.

338. Posner, J.; Polanczyk, G. V; Sonuga-Barke, E. Attention-deficit hyperactivity disorder. *Lancet* **2020**, *395*, 450–462, doi:10.1016/S0140-6736(19)33004-1.

339. Rohde, P.D.; Jensen, I.R.; Sarup, P.M.; Orsted, M.; Demontis, D.; Sorensen, P.; Kristensen, T.N. Genetic Signatures of Drug Response Variability in Drosophila melanogaster. *Genetics* **2019**, *213*, 633–650, doi:10.1534/genetics.119.302381.

340. Mackay, T.F.; Richards, S.; Stone, E.A.; Barbadilla, A.; Ayroles, J.F.; Zhu, D.; Casillas, S.; Han, Y.; Magwire, M.M.; Cridland, J.M.; et al. The Drosophila melanogaster Genetic Reference Panel. *Nature* **2012**, *482*, 173–178, doi:10.1038/nature10811.

341. Rochester, J.R.; Bolden, A.L.; Kwiatkowski, C.F. Prenatal exposure to bisphenol A and hyperactivity in children: a systematic review and meta-analysis. *Env. Int* **2018**, *114*, 343–356, doi:10.1016/j.envint.2017.12.028.

342. Crain, D.A.; Eriksen, M.; Iguchi, T.; Jobling, S.; Laufer, H.; LeBlanc, G.A.; Guillette, L.J. An ecological assessment of bisphenol-A: Evidence from comparative biology. *Reprod. Toxicol.* **2007**, *24*, 225–239, doi:10.1016/j.reprotox.2007.05.008.

343. Kaur, K.; Simon, A.F.; Chauhan, V.; Chauhan, A. Effect of bisphenol A on Drosophila melanogaster behavior--a new model for the studies on neurodevelopmental disorders. *Behav Brain Res* **2015**, *284*, 77–84, doi:10.1016/j.bbr.2015.02.001.

344. Richter, C.A.; Birnbaum, L.S.; Farabollini, F.; Newbold, R.R.; Rubin, B.S.; Talsness, C.E.; Vandenbergh, J.G.; Walser-Kuntz, D.R.; vom Saal, F.S. In vivo effects of bisphenol A in laboratory rodent studies. *Reprod Toxicol* **2007**, *24*, 199–224, doi:10.1016/j.reprotox.2007.06.004.

345. Rochester, J.R. Bisphenol A and human health: a review of the literature. *Reprod Toxicol* **2013**, *42*, 132–155, doi:10.1016/j.reprotox.2013.08.008.

346. Adesman, A.; Soled, D.; Rosen, L. Formula Feeding as a Risk Factor for Attention-Deficit/Hyperactivity Disorder: Is Bisphenol A Exposure a Smoking Gun? *J. Dev. Behav. Pediatr.* **2017**, *38*, 545–551, doi:10.1097/DBP.0000000000000468.

347. Braun, J.M.; Kalkbrenner, A.E.; Calafat, A.M.; Yolton, K.; Ye, X.; Dietrich, K.N.; Lanphear, B.P. Impact of early-life bisphenol A exposure on behavior and executive function in children. *Pediatrics* **2011**, *128*, 873–882, doi:10.1542/peds.2011-1335.

348. Hong, S.B.; Hong, Y.C.; Kim, J.W.; Park, E.J.; Shin, M.S.; Kim, B.N.; Yoo, H.J.; Cho, I.H.; Bhang, S.Y.; Cho, S.C. Bisphenol A in relation to behavior and learning of school-age children. *J Child Psychol Psychiatry* **2013**, *54*, 890–899, doi:10.1111/jcpp.12050.

349. Perera, F.; Vishnevetsky, J.; Herbstman, J.B.; Calafat, A.M.; Xiong, W.; Rauh, V.; Wang, S. Prenatal bisphenol a exposure and child behavior in an inner-city cohort. *Env. Heal. Perspect* **2012**, *120*, 1190–1194, doi:10.1289/ehp.1104492.

350. Yolton, K.; Xu, Y.; Strauss, D.; Altaye, M.; Calafat, A.M.; Khoury, J. Prenatal exposure to bisphenol A and phthalates and infant neurobehavior. *Neurotoxicol Teratol* **2011**, *33*, 558–566, doi:10.1016/j.ntt.2011.08.003.

351. Belfer, S.J.; Bashaw, A.G.; Perlis, M.L.; Kayser, M.S. A Drosophila model of sleep restriction therapy for insomnia. *Mol. Psychiatry* **2021**, *26*, 492–507, doi:10.1038/s41380-019-0376-6.

352. Miller, C.B.; Espie, C.A.; Epstein, D.R.; Friedman, L.; Morin, C.M.; Pigeon, W.R.; Spielman, A.J.; Kyle, S.D. The evidence base of sleep restriction therapy for treating insomnia disorder. *Sleep Med Rev* **2014**, *18*, 415–424, doi:10.1016/j.smrv.2014.01.006.

353. McBride, S.M.; Choi, C.H.; Wang, Y.; Liebelt, D.; Braunstein, E.; Ferreiro, D.; Sehgal, A.; Siwicki, K.K.; Dockendorff, T.C.; Nguyen, H.T.; et al. Pharmacological rescue of synaptic plasticity, courtship behavior, and mushroom body defects in a Drosophila model of fragile X syndrome. *Neuron* **2005**, *45*, 753–764.

354. Kochinke, K.; Zweier, C.; Nijhof, B.; Fenckova, M.; Cizek, P.; Honti, F.; Keerthikumar, S.; Oortveld, M.A.; Kleefstra, T.; Kramer, J.M.; et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual

Disability into Biologically Coherent Modules. *Am J Hum Genet* **2016**, *98*, 149–164, doi:10.1016/j.ajhg.2015.11.024.

355. Oortveld, M.A.; Keerthikumar, S.; Oti, M.; Nijhof, B.; Fernandes, A.C.; Kochinke, K.; Castells-Nobau, A.; van Engelen, E.; Ellenkamp, T.; Eshuis, L.; et al. Human intellectual disability genes form conserved functional modules in Drosophila. *PLoS Genet* **2013**, *9*, e1003911, doi:10.1371/journal.pgen.1003911.

356. Bellato, A.; Arora, I.; Hollis, C.; Groom, M.J. Is autonomic nervous system function atypical in attention deficit hyperactivity disorder (ADHD)? A systematic review of the evidence. *Neurosci. Biobehav. Rev.* **2020**, *108*, 182–206, doi:10.1016/j.neubiorev.2019.11.001.

357. Garcia-Rill, E. *Arousal in neurological and psychiatric diseases*; Elsevier: San Deigo, 2019; ISBN 9780128179925.

358. Cortese, S.; Moreira-Maia, C.R.; St Fleur, D.; Morcillo-Peñalver, C.; Rohde, L.A.; Faraone, S. V. Association between ADHD and obesity: A systematic review and meta-analysis. *Am. J. Psychiatry* **2016**, *173*, 34–43, doi:10.1176/appi.ajp.2015.15020266.

359. Fliers, E.A.; Buitelaar, J.K.; Maras, A.; Bul, K.; Höhle, E.; Faraone, S. V.; Franke, B.; Rommelse, N.N.J. ADHD is a risk factor for overweight and obesity in children. *J. Dev. Behav. Pediatr.* **2013**, *34*, 566–574, doi:10.1097/DBP.0b013e3182a50a67.

360. Miyazaki, C.; Koyama, M.; Ota, E.; Swa, T.; Mlunde, L.B.; Amiya, R.M.; Tachibana, Y.; Yamamoto-Hanada, K.; Mori, R. Allergic diseases in children with attention deficit hyperactivity disorder: a systematic review and meta-analysis. *BMC Psychiatry* **2017**, *17*, 120, doi:10.1186/s12888-017-1281-7.

361. Mota, N.R.; Poelmans, G.; Klein, M.; Torrico, B.; Fernandez-Castillo, N.; Cormand, B.; Reif, A.; Franke, B.; Arias Vasquez, A. Cross-disorder genetic analyses implicate dopaminergic signaling as a biological link between Attention-Deficit/Hyperactivity Disorder and obesity measures. *Neuropsychopharmacology* **2020**, doi:10.1038/s41386-019-0592-4.

362. Nielsen, P.R.; Benros, M.E.; Dalsgaard, S. Associations Between Autoimmune Diseases and Attention-Deficit/Hyperactivity Disorder: A Nationwide Study. *J Am Acad Child Adolesc Psychiatry* **2017**, *56*, 234-240 e1, doi:10.1016/j.jaac.2016.12.010.

363. Nigg, J.T.; Johnstone, J.M.; Musser, E.D.; Long, H.G.; Willoughby, M.T.; Shannon, J. Attention-deficit/hyperactivity disorder (ADHD) and being overweight/obesity: New data and meta-analysis. *Clin Psychol Rev* **2016**, *43*, 67–79, doi:10.1016/j.cpr.2015.11.005.

364. Tylee, D.S.; Sun, J.; Hess, J.L.; Tahir, M.A.; Sharma, E.; Malik, R.; Worrall, B.B.; Levine, A.J.; Martinson, J.J.; Nejentsev, S.; et al. Genetic correlations among psychiatric and immune-related phenotypes based on genome-wide association data. *Am J Med Genet B Neuropsychiatr Genet* **2018**, *177*, 641–657, doi:10.1002/ajmg.b.32652.

365. Caygill, E.E.; Brand, A.H. The GAL4 system: A versatile system for the manipulation and analysis of gene expression. *Methods Mol. Biol.* **2016**, *1478*, 33–52, doi:10.1007/978-1-4939-6371-3_2.

366. McGuire, S.E.; Mao, Z.; Davis, R.L. Spatiotemporal gene expression targeting with the TARGET and gene-switch systems in Drosophila. *Sci STKE* **2004**, *2004*, pl6, doi:10.1126/stke.2202004pl6.

367. Mohr, S.E.; Hu, Y.; Kim, K.; Housden, B.E.; Perrimon, N. Resources for functional genomics studies in Drosophila melanogaster. *Genetics* **2014**, *197*, 1–18, doi:10.1534/genetics.113.154344.

368. Schlegel, P.; Costa, M.; Jefferis, G.S. Learning from connectomics on the fly. *Curr Opin Insect Sci* **2017**, *24*, 96–105, doi:10.1016/j.cois.2017.09.011.

369. Chao, H.T.; Liu, L.; Bellen, H.J. Building dialogues between clinical and biomedical research through cross-species collaborations. *Semin. Cell Dev. Biol.* **2017**, *70*, 49–57, doi:10.1016/j.semcdb.2017.05.022.

370. Senturk, M.; Bellen, H.J. Genetic strategies to tackle neurological diseases in fruit flies. *Curr Opin Neurobiol* **2018**, *50*, 24–32, doi:10.1016/j.conb.2017.10.017.

371. Takano-Shimizu-Kouno, T.; Ohsako, T. Humanized Flies and Resources for Cross-Species Study. *Adv Exp Med Biol*

**2018**, *1076*, 277–288, doi:10.1007/978-981-13-0529-0_15.

372. Narayanan, A.S.; Rothenfluh, A. I Believe I Can Fly!: Use of Drosophila as a Model Organism in Neuropsychopharmacology Research. *Neuropsychopharmacology* **2016**, *41*, 1439–1446, doi:10.1038/npp.2015.322.

## SUPPLEMENTARY TABLES

**Supplementary Tables 1.** List of mouse strains with mutations in a single gene that present ADHD-related symptoms found in the Jackson database (only available online).

**Supplementary Table S2**. List of genes from mouse strains form the Jackson database that have been previously related with ADHD and/or its comorbid phenotypes

**Supplementary Table 3**. List of mouse strains with ADHD-related symptoms found in the Jackson database (only available online).

**Supplementary Table S2**. List of genes from mouse strains form the Jackson database that have been previously related with ADHD and/or its comorbid phenotypes

| Gene | Papers on ADHD | Disorder 1 | Papers 1 | Disorder 2 | Papers 2 | Disorder 3 | Papers 3 |
|---|---|---|---|---|---|---|---|
| *Anks1b* | 31388001 | ASD | SFARI | SCZ | 31250731 | | |
| *Atp1a3* | - | SUD | | anxiety | 22933743; 33326973; 20301294 | | |
| *Camk2a* | - | MD | - | Aggression | 26288127 | | |
| *Cdkl5* | 25315662 | ASD | SFARI | | | | |
| *Cntnap2* | 19546859; 19546859; 16571880; 23714751 | ASD | SFARI | | | | |
| *Crebbp* | - | ASD | SFARI | | | | |
| *Disc1* | 23389941; 28097908; 26649006 | ASD | SFARI | | | | |
| *Drd1* | 22404661; 28237458; 19695183 | SUD | 30268777; 29383684; 25907750; 19179847 | | | | |
| *Drd2* | 29984470; 22610946;18591481; 30946941; 30862909; | SUD | 32260442; 30118972; 26146874; 30268777; 19179847 | | | | |
| *Drd3* | 30051166 | SUD | 30268777; 25907750; 19179847 | | | | |
| *Dtnbp1* | - | SUD | 20615259 | | | | |
| *En2* | - | ASD | SFARI | | | | |
| *Esr1* | 28617822 | Aggression | 26288127 | | | | |
| *Fmr1* | 22101959; 20981777; 30483160 | ASD | SFARI | aggression | 24664669; 18570292 | anxiety | 30483160; 33195422 |
| *Fxr2* | - | SCZ | 18445270 | | | | |
| *Gabrb3* | - | ASD | SFARI | | | | |
| *Gria1* | - | ASD | SFARI | SUD | 29338492; 26101849; 27863698 | | |
| *Grin2b* | 17010153; 27818011 | ASD | SFARI | | | | |

| Gene | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Hmox1* | - | SCZ | 15098003; 25937794 | | | | | |
| *Htt* | - | ASD | SFARI | | | | | |
| *Il6* | 28262601; 32418647 | MD | 27555379; 30944309 | anxiety | 32769481; 30199144 | | | |
| *Mapt* | - | ASD | SFARI | | | | | |
| *Nlgn2* | 26152839 | ASD | SFARI | anxiety | 27865048 | | | |
| *Nlgn3* | - | ASD | SFARI | | | | | |
| *Npas3* | 21654738 | ASD | - | SCZ | 24674381; 20466522; 28499489 | | | |
| *Nr4a2* | 15635701 | MD | 23064081; 23087602 | SUD | - | | | |
| *Ppargc1a* | 31379624 | MD | 30381832; 21630437 | | | | | |
| *Rgs4* | 19124687 | ASD | - | | | | | |
| *Scn1a* | - | ASD | SFARI | | | | | |
| *Shank2* | - | ASD | SFARI | | | | | |
| *Shank3* | - | ASD | SFARI | SCZ | 20385823; 28371232 | SUD | - | |
| *Slc6a3* | 29984470;18591481 | SUD | 26146874; 31087723; 29332099 | anxiety | 28870407; 27343365; 24021960 | | | |
| *Syngap1* | 23161826 | ASD | SFARI | SCZ | 25034949; 30705251 | | | |
| *Uba6* | 26284580 | ASD | - | | | | | |