UNIVERSITAT DE BARCELONA

Caracterización clínica y molecular del síndrome de Rett: elucidar los casos no resueltos

Silvia Vidal Falcó

Aquesta tesi doctoral està subjecta a la llicència ***Reconeixement 4.0. Espanya de Creative Commons***.

Esta tesis doctoral está sujeta a la licencia ***Reconocimiento 4.0. España de Creative Commons***.

This doctoral thesis is licensed under the ***Creative Commons Attribution 4.0. Spain License***.

Caracterización clínica y molecular del síndrome de Rett: elucidar los casos no resueltos

Silvia Vidal Falcó

2020

Caracterización clínica y molecular del síndrome de Rett: elucidar los casos no resueltos

Memoria presentada por:
Silvia Vidal Falcó

Para optar al grado de
Doctora
Por la Universitat de Barcelona

Programa de genética humana, departamento de genética
Facultad de Biología

Tesi realizada bajo la dirección de la
Dra. Judith Armstrong Morón
en el Institut de Recerca Sant Joan de Déu

Judith Armstrong Morón	Daníel-Raul Grinberg Vaisman	Silvia Vidal Falcó
Directora	Tutor	Doctoranda

Barcelona, Mayo 2020

Als meus àngels

De com he aconseguit arribar fins aquí: Agraïments

Aquest es el resultat obtingut després d'un llarg camí ple d'emocions. Durant aquest anys he crescut com a persona, he après i millorat en tots els àmbits: a nivell acadèmic, professional i, sobretot, personal. He passat per moments molt diversos, des de alegria i eufòria a tensió i frustració en els que un no sap com seguir endavant, però, que amb temps, paciència i, sobretot, l'ajuda de tots els que han estat al meu costat he sigut capaç d'arribar als meus objectius. Per això, no voldria acabar aquesta etapa de la meva vida sense agrair-lis tot el seu recolzament i ajuda:

En primer lloc voldria agrair a l'*Insituto de Salud Carlos III* pel projecte FIS amb el qual hem pogut finançar gran part d'aquest projecte. També, i no menys important, als pacients i les seves famílies, juntament amb les associacions (Associació catalana de la síndrome de Rett, al Fondo Biorett y a Mi princesa Rett) que amb el seu gran esforç ens han impulsat a crear un grup de recerca dedicat a la síndrome de Rett cada cop més potent.

Als meus pares i la meva germana, voldria agrair i dedicar aquesta tesis a ells, son els principals impulsors de la meva persona, els fonaments del meu desenvolupament, els que des de petita m'han motivat a ser qui soc i gracies a ells soc capaç d'assolir totes les fites que m'imposo en aquesta vida.

Als meus nebots, Joan Enric i Pau, que son capaços d'alegrar-te el dia només amb un somriure i una abraçada.

A Dan, la ayuda que me has brindado ha sido sumamente importante, tanto a nivel intelectual como personal. Con tu seguridad y conocimiento has sido mi motor para seguir siempre adelante.

A mi apoyo en el laboratorio de investigación. Paola y a Laura, sin vuestra ayuda y nuestras sesiones de *retro-negativa* esto no hubiera sido lo mismo. Clara, Ainhoa y Gonzalo, sin vuestra apoyo y colaboración, esto no hubiera llegado tan lejos ni hubiéramos conseguido tanto como equipo Rett. Judith i Montse, per fer de mainaderes de les meves cèl·lules, cobrir-me en tot moment i ser un gran recolzament en moments complicats. ¡Tengo mucha suerte de teneros a todos vosotros!

A l'Edgar, la Nuri i la Dèlia, els qui em van introduir i guiar des de cero en un laboratori vinculat a un Hospital. La seva dedicació i esforç m'han ensenyat que cada tub no és només un simple número.

Al Jan i al Guerau, heu sigut essencials, gran part bioinformàtica d'aquesta tesis va ser molt complicada fins que veu arribar vosaltres.

A tot l'equip de Neurogenética y Medicina Molecular del Institut de Recerca Pediàtrica de l'Hospital Sant Joan de Déu, que em van acollir al seu laboratori com una més i m'han ensenyat en primera línia el que és la recerca bàsica.

En últim lloc, i no per això el menys important, voldria donar les gràcies a la meva directora de tesis, que ha estat al meu costat en tot moment y ha fet tot el possible per a que aquesta tesis arribi a bon fi. Judith, gràcies per brindar-me la oportunitat d'emprendre aquesta tesis i, sobretot, pel teu recolzament i per estar al meu costat, inclús en la distancia, en els moments més complicats d'aquesta etapa.

M'he fet més forta, i ja no sóc la mateixa que era quan vaig començar aquest camí. Només puc dir que tot això ha valgut la pena i sóc molt feliç de tenir-os a tots vosaltres en la meva vida.

Gràcies!

It is our choices that show what we truly are, far more than our abilities.

J.K. ROWLING, *Harry Potter and the Chamber of Secrets*

Abstract

Rett syndrome is a genetically based neurodevelopmental disorder that is included in the rare disease group due to its low incidence in the population. However, this disease is the second cause of severe intellectual disability in women after Down syndrome. This disorder was first described in 1966 by Dr. Andreas Rett, although it was not until 1963 that the Swedish pediatrician Hagberg defined it as Rett syndrome. Nevertheless, it was not until 20 years after the genetic cause behind the syndrome, the *MECP2* gene malfunction, was clarified. This finding allowed the genetic diagnosis in a large part of patients with a clinical diagnosis of RTT. However, approximately 5% of patients with classic RTT and more than 25% of patients with atypical forms present negative results for mutations in the *MEPC2* gene. This led to the search for other possible genes involved in RTT, especially in the case of atypical forms. Subsequent studies in search of a genetic cause for those patients without diagnosis made it possible to link the *CDKL5* and *FOXG1* genes to this pathology in 2004 and 2008, respectively. Even so, there is a percentage of patients with the Rett phenotype or similar that their ethology still remains unknown. Recently, next generation sequencing has promoted genetic diagnoses because of the quickness and afford ability of the method. Thus, the introduction of these new technologies allowed us to study a larger number of genes associated with RTT or similar phenotypes simultaneously, providing a genetic diagnosis for a wider group of patients. These new findings allow us to provide the clinician with more information and clues that could help in the prevention of future symptoms or in pharmacologic therapy. Likewise, these results have allowed us to understand the complexity of the disorder a little better.

ABREVIATURAS

AAV	Virus adeno-asociados
ACMG	American College of Medical Genetics
ADN	Ácido desoxiribonucléico
ARN	Ácido ribonucléico
CDKL5	*Cyclin-dependent-kinase-like 5,* gen en humanos
CNV	*Copy number variation*
CREB	cAMP *response element-binding*
CTD	Dominio C-terminal
dNTP	Desoxirribonucleótidos trifosfatos
DPR	*Deletion prone region*
FBD	*Forkhead binding domain*
FOXG1	*Forkhead box G1*, gen en humanos
FTY720	Fingolimod
HDAC	Histona desacetilasas
ICX	Inactivación del cromosoma X
IGF1	*Insulin-like growth factor-1*
iPSC	Células madre pluripotentes inducidas
Mb	Mega base
MBD	Dominio de unión a islas CpG metiladas
MECP2	*Methil CpG binding protein 2,* el gen en humano
MeCP2	*Methil CpG binding protein 2,* la proteína en humano
Mecp2	*Methil CpG binding protein 2,* la proteína en ratón
MeCP2-e1	Isoforma 1 de MeCP2
MeCP2-e2	Isoforma 2 de MeCP2
miARN	*Micro* ARN
NGS	*Next generation sequencing,* secuenciación masiva
NHGRI	*National Human Genome Research Institute*
NLS	Señal de localización nuclear
NTD	Dominio N-terminal
OMIM	*Online Mendelian Inheritance in Man*®
pb	Pares de bases
RTT	Síndrome de Rett
siARN	*small interfering* ARN
SNC	Sistema nervioso central
SNV	*Single nucleotide variant*
ssADN	ADN de cadena simple
TRD	Dominio de represión transcripcional
VUS	Variante de significado clínico incierto
WES	*Whole exome sequencing*
WGS	*Whole genome sequencing*

INTRODUCCIÓN

Síndrome de Rett

El síndrome de Rett (RTT; OMIM#312750) es un trastorno del neurodesarrollo de inicio precoz que afecta principalmente a niñas, con una incidencia de 1:10.000 niñas nacidas y la segunda causa de discapacidad intelectual en mujeres después del síndrome de Down. El RTT es una enfermedad de base genética, y hasta el momento, una gran cantidad de evidencias respaldan que variantes patogénicas en el gen *Methil CpG binding protein 2* (*MECP2*; OMIM*300005) son la causa principal que conducen al desarrollo de esta enfermedad. Este gen se encuentra localizado en el cromosoma X, presentando un patrón de herencia dominante ligado al X, por lo que hace que este síndrome afecte casi en su totalidad a niñas ya que los niños que presentan una variante patogénica en este gen fallecen antes de su nacimiento o de manera muy precoz.

Este síndrome fue descrito inicialmente por el pediatra Andreas Rett en el 1966 al observar a dos pacientes que presentaban el mismo fenotipo: microcefalia adquirida con atrofia cerebral, estereotipias de "lavado de manos" y pérdida de habilidades adquiridas después de un período de normalidad inicial (Rett, 1966). Sin embargo, no fue hasta el 1983 cuando el doctor Bengt Hagberg presentó a la comunidad médica una descripción clínica concisa de 35 pacientes acuñando el término de "*Síndrome de Rett*" para la enfermedad. Todas ellas, provenientes de tres países diferentes (Francia, Portugal y Suecia), presentaban una encefalopatía progresiva uniforme. Después de un desarrollo aparentemente normal durante los primeros 6-18 meses de vida, se producía un estancamiento seguido de una regresión de las funciones motoras y psíquicas adquiridas hasta el momento. En el transcurso de los meses el deterioro progresa a demencia severa, una pérdida de interés por el entorno, epilepsia, ataxia, microcefalia adquirida y pérdida del uso propositivo de las manos. Con el paso del tiempo, aparece un retraso en el crecimiento, a menudo acompañado con problemas de malnutrición y disfunciones respiratorias como apneas e hiperventilación. Finalmente, las pacientes llegan a un estado de aparente estabilidad de sus síntomas que puede llegar a durar décadas (Hagberg et al., 1983).

La esperanza de vida en las pacientes con RTT depende de la gravedad de las manifestaciones clínicas: presencia de epilepsia, infecciones recurrentes, la no deambulación autónoma, escoliosis y/o hiperlaxitud troncoencefálica. A pesar de esta larga lista de complicaciones tan prematuras, se estima que el 60% de las pacientes con RTT llegan a los

37 años de edad (Anderson et al., 2014), e incluso algunas pueden llegar a vivir hasta los 70 años (Hagberg et al., 2002).

1. Aspectos clínicos

RTT se encuentra dividido en dos categorías, la forma clásica y las formas atípicas. El criterio diagnóstico para el RTT es puramente clínico, no existe ningún biomarcador hasta la fecha que se identifique con el trastorno. Su diagnóstico se realiza en base a una serie de criterios clínicos definidos en la Tabla 1. Estos criterios fueron redactados por Hagberg y un grupo internacional de investigadores en 1988, que periódicamente son revisados y actualizados por neuropediatras expertos (2002 y 2010) (Hagberg et al., 2002, Neul et al., 2010).

Tabla 1. Criterios diagnósticos del RTT. *Modificado de Neul et al. 2010*

Criterios principales
Pérdida parcial o total de las habilidades manuales.
Pérdida parcial o total del lenguaje verbal.
Incapacidad para la marcha o marcha apráxica.
Estereotipias manuales (aleteo, lavado, palmas, retorcer manos, etc…).
Criterios de apoyo para la forma atípica
Alteraciones de la respiración en vigilia.
Bruxismo en vigilia.
Alteración del patrón de sueño.
Tono muscular anómalo.
Alteraciones de la vascularización periférica: manos y pies pequeños y fríos.
Escoliosis/cifosis.
Retraso de crecimiento.
Microcefalia adquirida.
Ataques de risa o gritos fuera de contexto.
Disminución de la respuesta al dolor.
Comunicación ocular intensa, o mejoría del contacto visual durante la evolución.
Criterios de exclusión
Daño cerebral secundario a trauma (peri o postnatal), enfermedad neurometabólica o infección grave que cause problemas neurológicos.
Desarrollo psicomotor claramente anómalo en los primeros 6 meses de vida.

Actualmente, estos criterios nos permiten definir a las pacientes RTT en dos formas diferenciadas: la forma típica o clásica y las formas atípicas o variantes del RTT (Tabla 2). La forma clásica es la expresión fenotípica más frecuente, descrita inicialmente por Rett y Hagberg, y definida mediante los cuatro criterios principales. Estos rasgos principales se desarrollan durante los primeros años de vida y permiten su diagnóstico entre los dos y cinco años de vida. Además, algunas características fenotípicas definidas como criterios de soporte pueden aparecer durante el transcurso de la enfermedad. Por otro lado, los pacientes diagnosticados con RTT atípico han de cumplir dos o más criterios principales más cinco o más criterios de apoyo para poder ser definidos como tal. Asimismo, han sido establecidos criterios que excluyen el diagnóstico de RTT: una lesión cerebral, una enfermedad neurometabólica, infección neurológica o un desarrollo psicomotor claramente anómalo con un inicio antes de los seis meses.

Tabla 2. Criterios para la distinción de las distintas formas del RTT. *Modificado de Neul et al. 2010.*

Requerimientos para la forma clásica
Un período de regresión seguido de una mejora o estabilización.
Cumplir todos los criterios principales y ninguno de exclusión.
Los criterios de soporto no son estrictamente necesarios que se cumplan.
Requerimientos para las formas atípicas
Al menos 2 de los 4 criterios principales.
5 de los criterios de soporte.

1.1 La forma clásica

Las pacientes diagnosticadas con la forma clásica del RTT presentan una disminución de la actividad voluntaria de las manos entre los 6 meses y los dos años de vida, pierden el lenguaje expresivo adquirido y el lenguaje receptivo se deteriora, con una gran afectación en el desarrollo psicomotor. Además, se asocia temporalmente una disfunción temporal de la comunicación y de socialización, incluyéndolas dentro de los trastornos del espectro autista. Una de las características más destacadas de este síndrome es la presencia de estereotipias manuales, en forma de "lavado de manos", retorciéndolas, ensalivándoselas, dando golpes y

automatismos de fricción. Durante el desarrollo motor, entre el primer y cuarto año de vida, empiezan a manifestarse anomalías en la deambulación, como la marcha apráxica, aunque hay pacientes que nunca llegan a adquirirla. El perímetro craneal es normal al nacer, pero entre los primeros cinco meses y los cuatro años de vida se observa una desaceleración del crecimiento cefálico, que en la mayoría de las pacientes evoluciona a una microcefalia adquirida.

La forma clásica tiene una progresión clínica típica, definida por 4 etapas:

- *Fase de normalidad*: donde se produce un desarrollo prenatal y perinatal aparentemente normal, aunque en el transcurso de las últimas décadas se ha descrito que durante este período se pueden apreciar ciertas anomalías difícilmente detectables. (Ver apartado *1.4 Diagnóstico anterior a la fase de regresión*).

- *Fase de regresión*: fase de duración variable entre las pacientes donde se detiene el desarrollo psicomotor que conduce a una pérdida profunda de habilidades adquiridas en las áreas de la comunicación, el uso de manos y la deambulación. Asimismo, en muchos casos también se enlentece el crecimiento cefálico desarrollando una microcefalia adquirida.

- *Fase de estabilización*: etapa pseudoestacionaria que transcurre durante los años preescolares y escolares en la que las pacientes muestran una amplia variedad de síntomas específicos de RTT.

- *Fase de declive*: en la última etapa de su vida, las pacientes experimentan un empeoramiento del rendimiento motor y disfunciones autónomas que se prolongan hasta el fallecimiento, lo que ocurre a edades extremadamente variables. Sin embargo, en aproximadamente una cuarta parte de los casos, las pacientes tienen una muerte súbita inesperada causada por inestabilidad cardíaca, crisis respiratoria o convulsiones no controladas (Chahrour and Zoghbi, 2007, Kyle et al., 2018).

1.2 Las formas atípicas

Existe una cohorte de pacientes que no cumplen con todos los criterios principales, pero comparten muchos rasgos fenotípicos con la forma clásica (aproximadamente el 20% de los casos). Estas pacientes con las variantes o formas atípicas pueden presentar un fenotipo tanto más grave, en el caso de las formas de epilepsia precoz y congénita, como mucho más leve, en el caso de la forma con lenguaje conservado y de regresión tardía:

- *Forma congénita o de inicio precoz*: es la forma más grave de RTT atípico y está caracterizada por un retraso global en el desarrollo psicomotor desde el nacimiento sin un periodo de aparente normalidad y sin fase de regresión. En estado más avanzado acaban apareciendo las estereotipias manuales y en la mayoría de los casos nunca se adquiere la deambulación (Rolando, 1985).

- *Epilepsia precoz*: presenta una evolución severa, el debut epiléptico ocurre entre las 6-10 semanas de vida con epilepsia parcial o generalizada resistente al tratamiento. El estancamiento del neurodesarrollo y las estereotipias se manifiestan posteriormente al inicio de las convulsiones (Hanefeld, 1985).

- *Forma con lenguaje conservado*: corresponde a pacientes que llegan a utilizar algunas palabras o frases cortas de forma propositiva. No llegan nunca a desarrollar un lenguaje oral normal, pero pueden llegar a mantener el lenguaje adquirido, o bien recuperarlo progresivamente durante la pubertad. En esta variante son poco frecuentes las crisis epilépticas y pueden no llegar a desarrollar microcefalia adquirida (Zappella, 1992).

- *Forma frustre o de regresión tardía,* en esta variante el periodo de regresión se produce entre el primer y tercer año de vida. Conservan parcialmente el uso de las manos y las estereotipias no son habituales, algunas veces sin detención del crecimiento cefálico. Las pacientes con regresión tardía mantienen más habilidades previamente adquiridas que las pacientes con la forma clásica y presentan un mejor pronóstico funcional (Hagberg and Skjeldal, 1994).

1.3 RTT en varones

Originalmente, la ausencia casi completa de varones con RTT clásico postulaba un efecto letal de las mutaciones en *MECP2* en ellos (Leonard et al., 2001, Tate et al., 1996). Por lo contrario, han sido reportadas e informadas mutaciones en *MECP2* en pacientes masculinos con una amplia variedad de presentaciones clínicas, incluidas discapacidades graves del desarrollo neurológico, encefalopatías congénitas y RTT clásico (Chahrour and Zoghbi, 2007). Los pacientes varones con sospecha de RTT o que presenten variantes patogénicas en el gen *MECP2* pueden clasificarse en varios grupos (Zeev et al., 2002):

- *Varones con síndrome de Klinefelter*: varones con dos copias del cromosoma X y una copia del cromosoma Y; por lo tanto, tienen un cromosoma adicional que les da una designación cromosómica de 47, XXY. Debido a que estos varones tienen dos copias del cromosoma X, se someten al mismo proceso de inactivación del X que se observa en las mujeres. Estos varones presentan una clínica muy consistente con la forma clásica. Además, presentan características del síndrome de Klinefelter, como el subdesarrollo de los genitales y la baja producción de hormonas sexuales.

- *Varones portadores de variantes patogénicas en mosaico en MECP2*: Dependiendo del porcentaje de mosaicismo celular, estos varones tendrán características clínicas similares a las mujeres con RTT clásico. Ha sido descrito que el 69% de los pacientes reportados con mosaicismo son varones (Zhang et al., 2019). De modo que el mosaicismo en *MECP2* debe tenerse en cuenta al diagnosticar clínicamente a niños con RTT.

- *Varones con encefalopatía neonatal severa asociada a MECP2*: Generalmente presentan un inicio más temprano y síntomas más graves que las mujeres. La mayoría de estos varones presentan problemas respiratorios, de alimentación y convulsiones de inicio temprano, pudiendo sobrevivir hasta la infancia tardía. Con frecuencia requieren un apoyo médico significativo, que incluye apoyo nutricional y respiratorio.

- *Casos esporádicos de varones que cumplen los criterios de inclusión para el RTT*: Varones con variantes patogéncias en *MECP2* fuera de las regiones del gen que generalmente se asocian con el RTT clásico en las mujeres y presentaciones más leves. Algunas de estas variantes, cuando se observan en mujeres, pueden presentarse asintomáticas o

con clínica muy leve. Se cree que estas mutaciones en ciertas regiones del gen *MECP2* no causan una alteración anómala significativa de la función de la proteína y, por lo tanto, producen síntomas más leves en mujeres y graves en varones (Chahrour and Zoghbi, 2007, Orrico et al., 2000).

1.4 Diagnóstico anterior a la fase de regresión

Debido al inicio fenotípico tardío de la enfermedad y la consecuente naturaleza centrada en la regresión de los criterios diagnósticos (Hagberg et al., 2002, Neul et al., 2010), la edad media de diagnóstico se encuentra alrededor de los 2 años para la forma clásica y un poco más de un año después para las formas atípicas, dificultando cualquier tipo de intervención temprana (Marschik et al., 2018, Tarquinio et al., 2015).

Aunque un desarrollo psicomotor aparentemente normal durante los primeros meses de vida se considera comúnmente una característica esencial y distintiva del RTT (Hagberg et al., 2002, Ricceri et al., 2008), las evidencias acumuladas durante las últimas décadas ponen en duda la ausencia de anomalías tempranas, por lo que se está cuestionando el concepto de regresión. En estudios retrospectivos se ha revelado que el 50% de los padres de las pacientes afectas pensaban que el desarrollo temprano de su hija había sido inusual. De forma recurrente los padres informaban que su bebé tenía la mirada "vacía", particularmente plácida, y dormía tan profundamente que tenían que despertarla la gran mayoría de veces para su alimentación (Witt-Engerstrom and Gillberg, 1987, Kerr et al., 1997, Naidu, 1997, Leonard and Bower, 1998). Dichos informes, aunque posiblemente sesgados por el recuerdo, proporcionaron la primera sugerencia de que el desarrollo temprano de las funciones psicomotoras, a pesar de parecer aparentemente normales, ya se muestra algunas pequeñas alteraciones. Las evaluaciones objetivas posteriores llevadas a cabo en las últimas dos o tres décadas por investigadores y médicos respaldan tales evidencias, desafiando así el paradigma del desarrollo temprano normal (Figura 1) (Kerr, 1995, Pokorny et al., 2018).

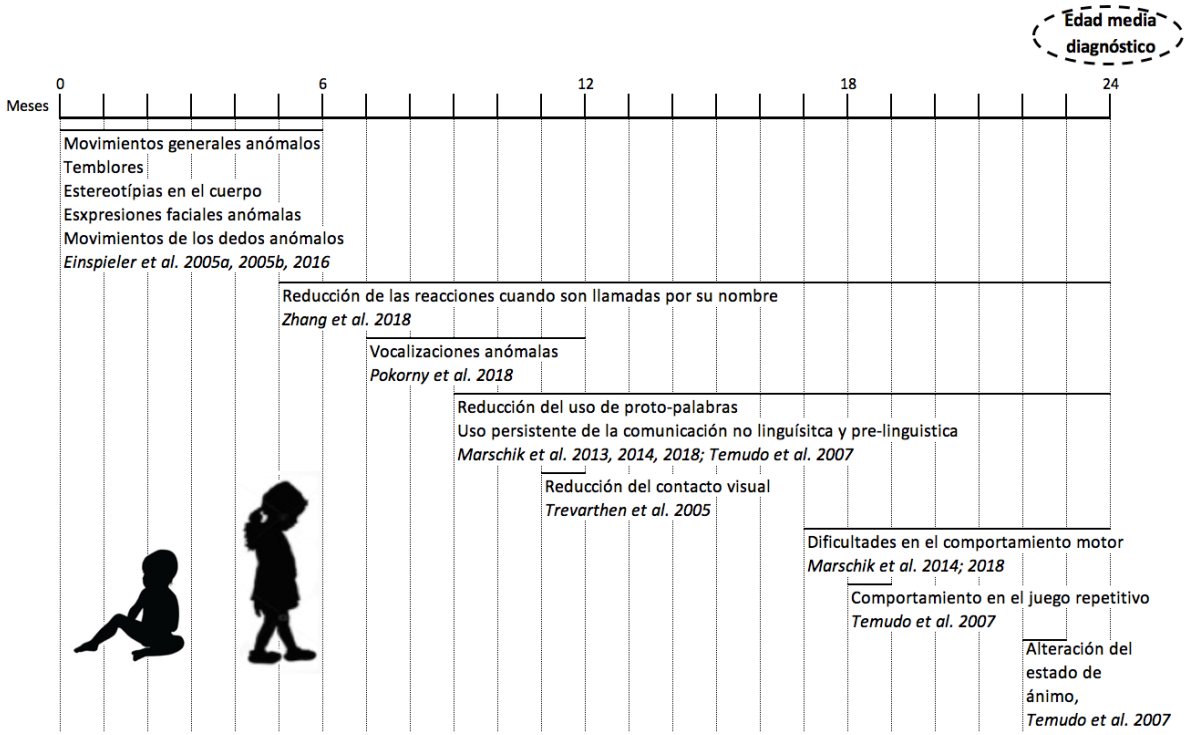


Figura 1. Alteraciones tempranas del comportamiento en pacientes con RTT. Visión general de las anomalías del comportamiento identificadas hasta ahora en los recién nacidos con RTT durante la fase de normalidad. La edad de diagnóstico más para RTT es de aproximadamente a los 2 años. *Imagen modificada de Cosentino et al. 2019.*

La revisión exhaustiva de Einspieler y Marschik (2019) resume estas características previas a la regresión. La mayor parte de este trabajo se ha basado en la adquisición de videos familiares durante el período de pre-regresión de pacientes diagnosticadas en última instancia con RTT. Aunque esta metodología tiene limitaciones, la naturaleza esporádica de la enfermedad requiere este enfoque. Las características previas a la regresión incluyen una desviación en la progresión de los "movimientos generales" y retrasos en el desarrollo a la presencia de características anómalas tales como estereotipias intermitentes de las manos o vocalizaciones cualitativamente anómalas. Aunque estas características de la vida temprana son interesantes para la creación de un algoritmo para identificar a los infantes afectados antes de la regresión, aún existen limitaciones. Primero, ninguna de estas características se encuentra de manera uniforme en todas las pacientes que terminan desarrollando RTT. Por ejemplo, solo el 34% presentaba estereotipias manuales tempranas. En segundo lugar, muchas de estas características requieren la determinación de expertos para reconocer estas

anomalías. Muchas no serían recogidas en evaluaciones clínicas de rutina o reconocidas por los padres, especialmente dada su naturaleza intermitente.

1.5 Heterogeneidad fenotípica del RTT

Dado el amplio rango de fenotipos y mutaciones que existen asociadas al RTT, surge la pregunta de qué tan bien se correlaciona el genotipo con el desarrollo fenotípico. Esta pregunta debe considerarse en dos niveles:

1. Dentro de las limitaciones de un diagnóstico clínico de RTT clásico o atípico, ¿ciertas mutaciones tienen más probabilidades de conducir a un cuadro clínico específico o una presentación general más grave?

2. En la amplia gama de fenotipos relacionados con el gen *MECP2*, ¿ciertas mutaciones son más propensas a conducir al RTT frente a otro trastorno ya descrito?

Se han realizado numerosos estudios de correlación genotipo-fenotipo para determinar cómo correlacionan grupos de mutaciones con la gravedad general del fenotipo de las pacientes, generalmente examinando pacientes con un diagnóstico de RTT clásico o atípico. Sin embargo, muchos de estos estudios no han podido establecer una correlación significativa entre la cohorte de pacientes estudiadas (Ham et al., 2005, Guerrini and Parrini, 2012). No obstante, algunos estudios han podido demostrar ciertas evidencias de correlación, centrándose en el efecto de las mutaciones y en la gravedad general de la presentación clínica en las pacientes con RTT (Cuddapah et al., 2014, Neul et al., 2008). Ha sido descrito que las mutaciones que truncan la proteína de forma prematura como p.R168*, p.R255* y p.R270* y las grandes inserciones y deleciones que crean un efecto deletéreo en la función de la proteína, causan un fenotipo más grave. Por otro lado, las mutaciones *missense* como p.R133C y p.R306C, las mutaciones que truncan la proteína más cerca de la región C-terminal, como las de p.R294* y otras en el extremo 3', están asociadas con fenotipos más leves. A pesar de esto, las variaciones de fenotipo ocurren comúnmente entre pacientes con la misma mutación. Se han propuesto algunas posibles explicaciones, como las diferencias en la inactivación del cromosoma X (Knudsen et al., 2006, Ehrhart et al., 2018) y la presencia de

una segunda mutación genética, que pueda actuar como modificador, empeorando o mejorando el cuadro clínico que acaban desarrollando las pacientes (Zeev et al., 2002).

Sin embargo, las comparaciones entre estudios son complicadas por una serie de factores, como la variabilidad en las herramientas de evaluación utilizadas para determinar la gravedad del fenotipo. Combinado con el hecho de que tanto en el RTT clásico como en el atípico, la edad de los sujetos juega un papel importante en la presentación de los síntomas, como la escoliosis es más frecuente en pacientes de edad avanzada, y que existe una interrelación entre varios componentes del fenotipo, como las pacientes sin deambulación tienen más probabilidades de tener escoliosis. Determinar con claridad las correlaciones específicas de genotipo-fenotipo no es una tarea trivial. De este modo, la existencia de casos con la misma mutación descrita, pero con diferencias en la gravedad clínica enfatiza la necesidad de proceder con precaución a la hora de predecir la posible progresión futura de las manifestaciones clínicas que pueda desarrolla una paciente en concreto.

Actualmente, en RTT, como en otras enfermedades, el término "*like*" se emplea en pacientes que no cumplen los criterios clínicos establecidos, pero presentan un fenotipo que concuerda con la enfermedad en mucho de los aspectos, como el retraso en el desarrollo psicomotor (con o sin regresión), movimientos estereotipados, pérdida del uso de las manos y del lenguaje. Hasta el momento, no se ha establecido y publicado un consenso de criterios formales para el diagnóstico de estos pacientes que presentan el síndrome de Rett-*like* (RTT-*like*) (Schonewolf-Greulich et al., 2019).

2. Aspectos genéticos

Las bases genéticas y el tipo de herencia del RTT fueron desconocidos durante décadas, desde que el doctor A. Rett (1966) y B. Hagberg (1983) la describieron. No fue hasta el año 1999, cuando se identificaron en pacientes afectas variantes patogénicas *de novo* en el gen *MECP2* (Amir et al., 1999), por lo que se verificó la hipótesis de que el RTT es una enfermedad ligada al cromosoma X dominante.

2.1 El gen *MECP2*

El gen *MECP2* se encuentra localizado en el brazo largo del cromosoma X (Xq28) flanqueado por los genes *RCP* e *IRAK1* y está formado por cuatro exones (Figura 2A y 2B). Este gen codifica para una proteína de unión al ADN metilado y actúa como factor regulador de la transcripción (MeCP2) que tanto puede inhibir como activar la transcripción de otros genes. Esta proteína es esencial para la maduración neuronal y el correcto funcionamiento de las células del sistema nervioso central (SNC) (Yasui et al., 2014).

MeCP2 tiene dos isoformas descritas (MeCP2-e1 y MeCP2-e2) con un *splicing* diferencial en los exones 1 y 2 que contribuyen a las diversas funciones de la proteína. MeCP2-e1 contiene los exones 1, 3 y 4, y el codón de inicio se encuentra en el exón 1, mientras que MeCP2-e2 contiene los exones 2, 3 y 4 y el codón de inicio se encuentra en el exón 2 (Figura 2C) (Ariani et al., 2008). Ambas isoformas contienen todos los dominios funcionales de la proteína. Sin embargo, MeCP2-e1 se expresa principalmente en el SNC (Kalscheuer et al., 2003), lo que sugiere que es la isoforma predominante en el cerebro, mientras que la isoforma MeCP2-e2 tiene niveles de expresión más elevados en el músculo esquelético, placenta, hígado y glándula prostática (Liyanage and Rastegar, 2014).

Actualmente, en la base de datos de la *NCBI Reference Sequence*, donde se proporcionan secuencias de referencia afianzadas para la anotación genómica, la identificación y caracterización de genes, han sido anotadas otras posibles candidatas a isoformas de MeCP2 con *splicing* alternativo y otros posibles exones, que a día de hoy no sabemos si pueden presentar variantes patogénicas que puedan estar implicadas en el desarrollo del RTT (https://wwwncbi.nlm.nih.gov/gene/4204).

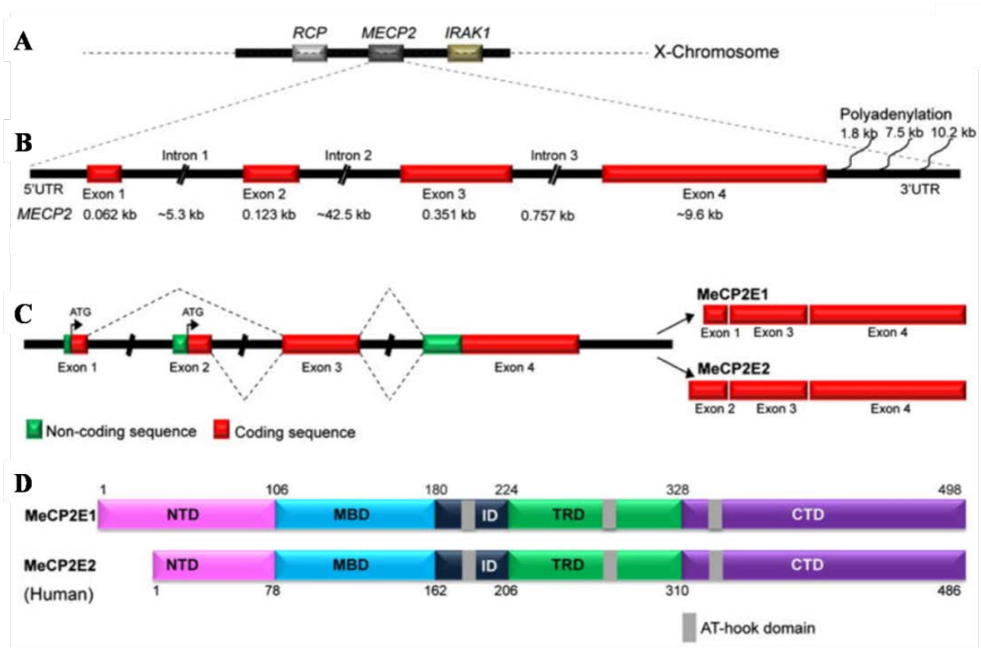


Figura 2. Estructura del gen *MECP2* y la proteína MeCP2. (a) El gen *MECP2* está ubicado en el cromosoma X (Xq28), flanqueado por los genes *RCP* e *IRAK*. (b) Esquematización del gen *MECP2,* compuesto por cuatro exones y tres intrones. El gen tiene tres sitios de poliadenilación en la región 3'UTR. Se indican los tamaños de cada región exónica e intrónica. (c) La generación de dos isoformas MeCP2: MeCP2E1 y MeCP2E2. El sitio de inicio de traducción (ATG) para cada isoforma se indica mediante flechas. Regiones codificantes marcada en rojo. Regiones no codificantes marcadas en verde. (d) Estructura proteica de MeCP2E1 (498 aminoácidos) y MeCP2E2 (486 aminoácidos). NTD, dominio N-terminal; MBD, dominio de unión a islas CpG metiladas; ID, región entre dominios; TRD, dominio de represión transcripcional; CTD, dominio C-terminal; AT-hook dominio localizado dentro del dominio TRD que permite la unión a regiones ricas en adenina-timina (AT). *Imagen modificada de Liyanage and Rastegar, 2014.*

MeCP2 presenta 2 dominios funcionales relevantes: el dominio MBD (*Methyl-CpG Binding Domain*) situado en la región N-Terminal y el dominio TRD (*Transcriptional Repressor Domain*) en la región central de la proteína. A parte, también presenta un dominio de unión con motivos WW localizados en la región C-terminal involucrado en la interacción proteína-proteína y dos señales de localización nuclear o NLS (*Nuclear Localization Signal*) (Figura 2D).

El dominio MBD se une exclusivamente a islas CpG metiladas simétricamente (Nan and Bird, 2001), aunque con menor preferencia también puede unirse al ADN no metilado (Meehan et al., 1992, Ballestar et al., 2000). El tipo de metilación a la que se une también es importante, se ha demostrado que MeCP2 es la principal proteína de unión a la 5-

hidroximetilcitosina en el cerebro, una marca epigenética enriquecida en genes neuronales activos (Mellen et al., 2012). La unión que ofrece el dominio MBD es importante para las funciones de MeCP2 (Lewis et al., 1992, Zhao et al., 2015). Esta unión puede activar o reprimir la transcripción de genes, dependiendo de los cofactores reclutados (Lewis et al., 1992, Chahrour et al., 2008, Yasui et al., 2007).

A diferencia del dominio MBD, el TRD no está todavía bien determinado, con estudios contradictorios relacionados con su supuesta función (Zlatanova, 2005). Principalmente, se cree que el dominio TRD es necesario para el papel silenciador transcripcional de la proteína, al unirse a proteínas co-represoras como mSin3A, lo que afecta el reclutamiento de histona desacetilasas y produce cambios en la arquitectura de la cromatina (Jones et al., 1998). Sin embargo, esta interacción es de baja frecuencia e inestable en estudios realizados *in vivo* (Klose and Bird, 2004).

2.1.1 Variantes patogénicas

La mayoría de las variantes patogénicas asociadas al RTT ocurren en los exones 3 y 4 del gen. Y, aunque también han sido detectadas en el exón 1, hasta el momento no se han identificado variantes patogénicas en el exón 2 en pacientes con RTT (Kalscheuer et al., 2003). Aunque existen ciertas regiones del gen descritas como puntos calientes para la aparición de mutaciones en el gen (*hotspots)*, hasta el momento, más de 800 variantes patogénicas han sido identificadas en el 95% de pacientes con RTT clásico y en el 75% de las pacientes con RTT atípico (Christodoulou et al., 2003, Neul et al., 2008). La mayoría de las variantes patogénicas se encuentran dentro del dominio TRD (34%) y el dominio MBD (25%). Ocho mutaciones principales representan aproximadamente el 47% de todas las mutaciones en *MECP2*. Estas mutaciones son p.R106T, p.R133C, p.T158M, p.R168*, p.R255*, p.R270*, p.R294* y p.R306C, que se distribuyen entre el dominio MBD y TRD (Figura 3). Además, el 98% de las mutaciones que se localizan en el dominio de la señal de localización nuclear (NLS) se consideran patogénicas. Del mismo modo, en una alta proporción de variantes observadas en la TRD (88%) y MBD (84%) se les ha asignado un estado patogénico, mientras que solo las variaciones en la región C-terminal (52%) se notan como patogénicas (Krishnaraj et al., 2017).

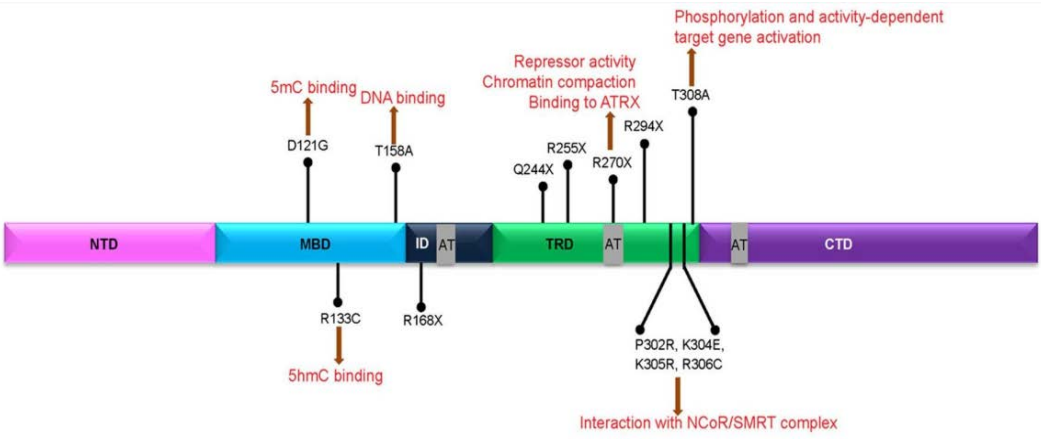


Figura 3. Distribución de variantes patogénicas conocidas en MeCP2 y las interacciones/funciones abolidas debido a ellas. Se ilustran las mutaciones bien conocidas que se encuentran en pacientes con síndrome de Rett y la pérdida de interacciones y funciones de MeCP2 debido a estas mutaciones. NTD, dominio N-terminal; MBD, dominio de unión a islas CpG metiladas; ID, región entre dominios; TRD, dominio de represión transcripcional; CTD, dominio C-terminal; AT, dominio AT; 5mC, 5-metilcitosina; 5hmC, 5-hidroximetilcitosina. Los números de aminoácidos están de acuerdo con la ubicación en la isoforma MeCP2E2. *Imagen obtenida de Liyanage and Rastegar, 2014.*

Aunque las mutaciones *MECP2* se consideran la causa más frecuente asociada al RTT y existe una fuerte correlación entre las mutaciones en el gen y los fenotipos RTT, se encuentran varios casos de pacientes RTT sin mutaciones *MECP2* (Hoffbuhr et al., 2001, Huppke and Gartner, 2005, Temudo et al., 2011). Asimismo, se han descrito pacientes que no muestran ninguno de los fenotipos RTT con mutaciones conocidas en *MECP2* ya claramente descritas como causantes de RTT (Suter et al., 2014). Además de RTT, se han encontrado mutaciones en *MECP2* asociadas con trastornos neurológicos y neuropsiquiátricos tales como esquizofrenia (Cohen et al., 2002), trastornos del espectro alcohólico fetal (Zoll et al., 2004), síndrome de psicosis, signos piramidales y macroorquidia (Klauck et al., 2002), trastorno del espectro autista (Beyer et al., 2002), síndrome de Prader-Willi (Tejada et al., 2006) y síndrome de Angelman (Watson et al., 2001). Estos estudios enfatizan que el diagnóstico de RTT, así como otros trastornos neurológicos relacionados con MeCP2, no deben limitarse a la detección de mutaciones en el gen *MECP2*, sino que debe llevarse a cabo junto con un exhaustivo reconocimiento clínico. También destacan la importancia de comprender la función de las mutaciones así como en otros trastornos asociados a MeCP2 (Liyanage and Rastegar, 2014).

2.1.2 Expresión de MeCP2

Más de dos décadas de investigación básica ha generado grandes avances en la comprensión de las funciones de MeCP2, así como de la fisiopatología del RTT. Elucidar las funciones de un gen en particular es primordial para comprender cómo su disfunción contribuye a la enfermedad. La patogenicidad molecular de las mutaciones en *MECP2* es compleja ya que están involucrados múltiples tejidos y funciones, y, en su mal funcionamiento, una gran cantidad de genes exhiben una expresión anómala con implicaciones en el equilibrio entre la excitación y la inhibición sináptica (Kron et al., 2012). Además de las mutaciones de pérdida de función, las duplicaciones en el gen también deterioran el SNC y conducen al desarrollo del síndrome de duplicación de *MECP2*, trastorno neurológico grave caracterizado por discapacidad intelectual y motora graves (Lombardi et al., 2015). De este modo, las drásticas consecuencias neurológicas tanto de la ganancia como la pérdida de función de *MECP2* remarcan la necesidad de comprender los roles de MeCP2 en el SNC.

La proteína MeCP2 tiene una expresión ubicua. Si bien sus funciones periféricas han sido descritas recientemente en la fatiga del ejercicio y las propiedades óseas (Ross et al., 2016), su acción principal y donde las concentraciones de la proteína son más elevadas es en el cerebro, lo que explica que la afectación principal en las pacientes es el en SNC. Además, los niveles de MeCP2 fluctúan durante el desarrollo del sistema nervioso donde son más elevados durante la etapa embrionaria, disminuyen al nacimiento y posteriormente alcanza su punto máximo durante el desarrollo sináptico y de la plasticidad (Figura 4) (Gos, 2013, Olson et al., 2014). Específicamente, los niveles de MeCP2 en las neuronas y los astrocitos aumentan durante el desarrollo postnatal: determinan la maduración, el tamaño neuronal y la ramificación dendrítica, hechos que se relacionan con la microcefalia adquirida y el comportamiento específico de los pacientes con RTT (Zoghbi, 2005, Chapleau et al., 2009, Mellios et al., 2018). Estos estudios sugirieron un papel primordial de MeCP2 en las últimas etapas del desarrollo neurológico y una función en las neuronas maduras y, además de una función prescindible de MeCP2 durante las primeras etapas embrionarias. De ahí que, se cree que este patrón de expresión tardía de MeCP2 es el que produce el desarrollo temprano aparentemente normal de las pacientes RTT. No obstante, un estudio más reciente demuestra que la expresión de MeCP2 se puede detectar en neuronas inmaduras de la corteza prenatal en ratones (Bedogni et al., 2016) y en células neuronales inducidas de células madre pluripotentes (iPSC) de humanos (Kim et al., 2011).

Asimismo, estudios clínicos han revelado alteraciones neurológicas en pacientes con RTT en los primeros meses de vida (Einspieler et al., 2005, Marschik et al., 2018). Por lo tanto, la expresión de MeCP2 desde las primeras etapas de desarrollo hasta la edad adulta tanto en ratones como en humanos y la presencia de déficits neurológicos en pacientes con RTT desde los primeros años de vida sugieren que MeCP2 regula todas las etapas del desarrollo neurológico y la función cerebral adulta (Gulmez Karaca et al., 2019). Sin embargo, los síntomas no neuronales asociados al RTT como las anomalías respiratorias, los problemas cardíacos, la escoliosis y los movimientos de las extremidades indican la importancia de la proteína fuera del SNC (Nomura and Segawa, 1992, Guideri and Acampa, 2005, Ogier and Katz, 2008)

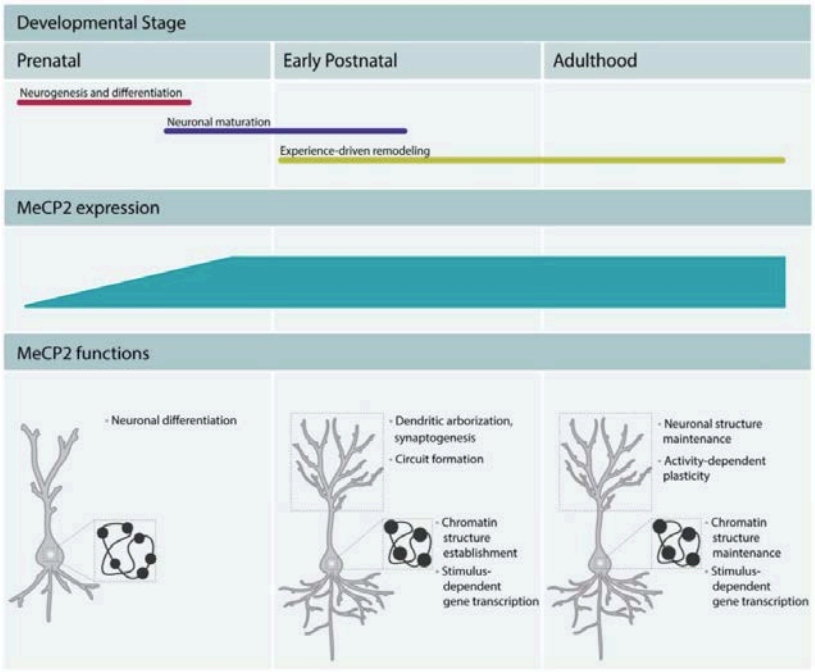


Figura 4. MeCP2 regula el desarrollo del cerebro y mantiene la función de las neuronas maduras durante la edad adulta. MeCP2 regula la diferenciación neuronal en el desarrollo embrionario temprano, la maduración neuronal y la formación de circuitos. MeCP2 promueve la agrupación de cromocentros durante la diferenciación y la maduración, y por lo tanto participa en el establecimiento de la estructura de cromatina típica de las neuronas maduras (las neuronas maduras presentan menos y más densos cromocentros, representados en la Figura como puntos negros en el núcleo). En la edad adulta, MeCP2 es un factor crítico en el mantenimiento de la función neuronal. Mantiene la estructura de la cromatina y regula el perfil transcriptómico neuronal. Además, parece mantener un estado permisivo para la transcripción del gen dependiente del estímulo y regular la función cognitiva. *Imagen obtenida de Gulmez Karaca et al., 2019.*

2.1.2 Funciones de MeCP2

MeCP2 es una proteína multifuncional que participa en la regulación transcripcional, así como en la modulación de la estructura de la cromatina. Los diferentes dominios de MeCP2 se han asignado para facilitar múltiples funciones a través de la unión directa del ADN, la interacción con otras proteínas o mediante el reclutamiento de otros factores (Guy et al., 2011).

El papel de MeCP2 en la regulación transcripcional se ha descrito en varios aspectos de la regulación génica, tanto en la represión como la activación transcripcional. Por ejemplo, la asociación de MeCP2 con complejos represores que contienen SIN3a e histona desacetilasas (HDAC) conduce a la represión transcripcional. Por otro lado, la activación transcripcional de genes mediada por MeCP2 ocurre en asociación con los complejos activadores que contienen la proteína de unión al elemento de respuesta a AMPc (CREB) (Chahrour et al., 2008). El reclutamiento de MeCP2 a los loci genómicos activos enriquecidos con 5 hmC (Mellen et al., 2012) y su asociación con la proteína TET1 (Cartron et al., 2013) respaldan aún más el papel de MeCP2 en la activación génica. Aunque se han reportado numerosos genes diana para MeCP2 en múltiples sistemas celulares (Zachariah and Rastegar, 2012), la naturaleza diversa de los genes diana de MeCP2 y los efectos opuestos en los genes estudiados plantean preguntas sobre la función exacta de MeCP2 como un regulador transcripcional.

Hay estudios que definen a MeCP2 como un modulador epigenético de todo el genoma en lugar de un regulador transcripcional (Della Ragione et al. 2012). Sin embargo, hasta la fecha, sigue siendo difícil saber cómo una sola proteína puede modular una gran variedad de funciones opuestas. Como proteína importante de unión al ADN, también se ha descrito que MeCP2 participa en el control de la estructura de la cromatina (Liyanage et al. 2012; Zlatanova 2005; Chadwick y Wade 2007). MeCP2 puede modular la arquitectura de la cromatina al condensar el ADN mediante la regulación de las interacciones de largo alcance (Horike et al. 2005), la formación de estructuras de cromatina de orden superior (Georgel et al. 2003; Agarwal et al. 2011) y la formación de bucles de cromatina y Puentes de ADN (Georgel et al. 2003; Nikitina et al. 2007b; Yasui et al. 2007). Además, hay estudios de localización de MeCP2 en regiones de heterocromatina teñidas con DAPI (Zachariah et al. 2012; Craig et al. 2003) y la agrupación de cromocentros durante la sobreexpresión de MeCP2 (Brero et al. 2005) han generado más evidencias adicionales sobre el papel de MeCP2

como una proteína arquitectónica de cromatina. Además, se han observado diferencias significativas en los tamaños y números de cromocentros en las neuronas Mecp2^{knockout} y Mecp2^{wildtype}, lo que respalda aún más el papel de MeCP2 en la organización de la cromatina (Singleton et al. 2011). Más importante aún, se ha demostrado que las mutaciones patogénicas en *MECP2* que causan RTT interrumpen la formación de estructuras de cromatina de orden superior (Nikitina et al. 2007a; Agarwal et al. 2011; Kumar et al. 2008).

2.2 Otros genes asociados al RTT: *CDKL5* y *FOXG1*

La identificación del gen *MECP2* y la mayor precisión de las técnicas diagnósticas han permitido llegar a un diagnóstico genético en gran parte de las pacientes con diagnóstico clínico de RTT. Sin embargo, aproximadamente el 5% de las pacientes con RTT clásico y más del 25% de las pacientes con formas atípicas tienen un estudio negativo para mutaciones en el gen *MEPC2*. Lo que llevó a la búsqueda de otros posibles genes implicados en el RTT, especialmente en el caso de las formas atípicas (Weaving et al., 2005).

2.2.1 Gen CDKL5

En el 2004 se describió el gen *Cyclin-dependent-kinase-like 5* (*CDKL5*; OMIM*300203) asociado con la forma atípica de epilepsia precoz (Tao et al., 2004). En los primeros estudios se proporcionó diagnóstico genético a un 3% de las pacientes sin diagnóstico molecular previo (Weaving et al., 2004). Este gen sigue un patrón de herencia ligada al X dominante ya que las mutaciones patogénicas descritas en las pacientes eran *de novo*, igual que en *MECP2*. Variantes patogénicas en este gen se han descrito también en niños con encefalopatía epiléptica resistente a fármacos (Mirzaa et al., 2013) o con autismo sin epilepsia, pero especialmente en la variante atípica de RTT con epilepsia precoz (Russo et al., 2009).

El gen *CDKL5* se encuentra localizado en el brazo corto del cromosoma X (Xp22), está formado por 21 exones y codifica para una proteína quinasa que participa en la migración y maduración neuronal (Figura 5A). Estudios experimentales sugieren que este gen es capaz de modular la expresión de la proteína MeCP2 y participar en vías comunes en ambos genes (Mari et al., 2005). En vista de la posibilidad de que *CDKL5* esté modulando la expresión de *MECP2* y del espectro fenotípico compartido que presentan las pacientes con variantes patogénicas en estos dos genes, se ha sugerido que ambos genes pueden desempeñar un papel común en un proceso de patogenicidad (Tao et al., 2004).

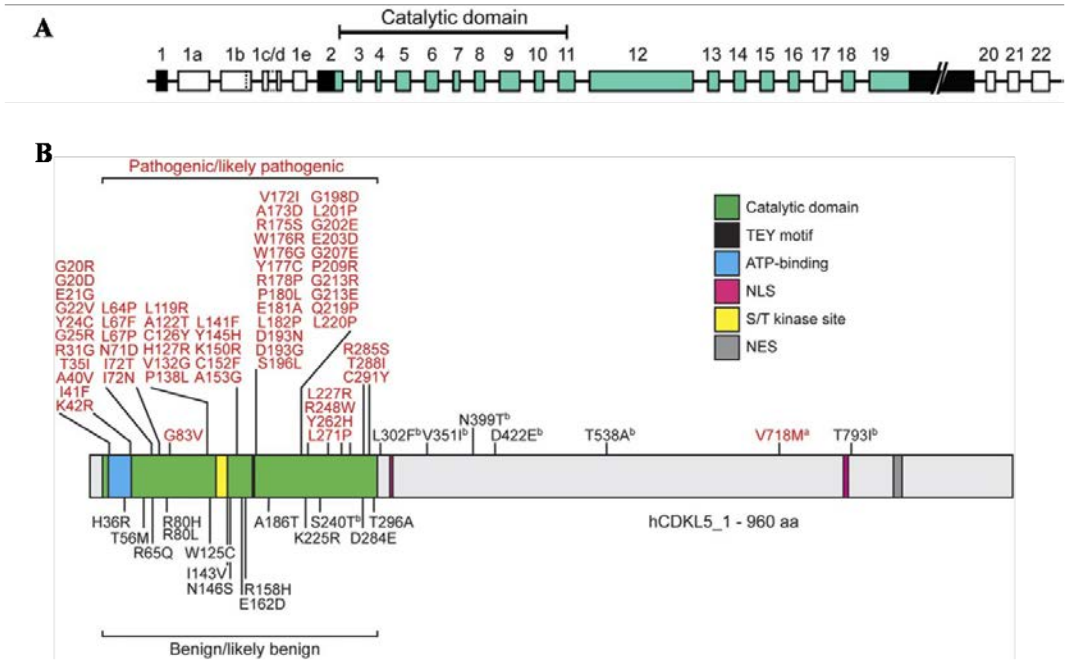


Figura 5. Estructura y distribución de mutaciones *missense* en el gen *CDKL5*. (A) Estructura del gen con la isoforma dominante en el cerebro, hCDKL5_1. De color azul-verde (regiones codificantes) y negro (UTR). Los intrones no están a escala. (B) Variantes *missense* patogénicas en *CDKL5*. Los dominios funcionales en la proteína CDKL5 están codificados por colores. Los números se refieren a las posiciones de los aminoácidos. Las variantes en rojo (superior) son patógenas o probablemente patógenas. Las variantes en negro (inferior) son benignas o probablemente benignas. a = variante con posible alteración en el *splicing*. b = variante de significado clínico incierto. NES = señal de exportación nuclear; NLS = señal de localización nuclear; ST = sitio activo de serina-treonina quinasa; TEY = motivo Thr-Glu-Tyr conservado. *Imagen obtenida de Hector et al., 2017*.

La patogenicidad de las variantes descritas en *CDKL5* debe ser probada en modelos animales o en líneas celulares humanas diseñadas (Wang et al., 2012). Sin embargo, el estudio realizado por Hector et al. 2017 proporciona evidencias acerca de la clasificación de variantes específicas en el gen e información para el diagnóstico genético (Hector et al., 2017). Aunque las variantes patogénicas de *CDKL5* se encuentran en la mayoría de las regiones codificantes del gen, las variantes *missense* claramente se agrupan en el dominio catalítico N-terminal. Es probable que las variantes *missense* fuera de este dominio y todas las variantes en los exones 20, 21 y 22 sean benignas (Figura 5B). En comparación con los síndromes de duplicación bien descritos que involucran genes estrechamente relacionados, *MECP2* y *FOXG1* (Van Esch et al., 2005, Brunetti-Pierri et al., 2011), se requieren más evidencias para concluir que existe un síndrome de duplicación de *CDKL5* bien definido.

2.2.2 Gen FOXG1

Cuatro años más tarde, en 2008, el gen *Forkhead box G1* (*FOXG1*; OMIM*164874), fue asociado a la forma congénita del RTT (Ariani et al., 2008). Ariani et al. 2008 identificaron dos variantes que truncaban la proteína en el gen, en dos pacientes con microcefalia, discapacidad intelectual y movimientos estereotipados similares a los que se observan en pacientes con RTT clásico. Siguiendo un patrón de herencia autosómica dominante, por lo que es común encontrar varones afectos que presentan la variante congénita del RTT atípico (Van der Aa et al., 2011). Los signos específicos de esta variante incluyen la presencia de un trastorno del movimiento discinético-hipercinético, la falta de regresión o arritmia respiratoria, y la aparición de malformaciones cerebrales en pacientes con una variante (Kortum et al., 2011). Las principales características clínicas observadas en asociación con variantes en *FOXG1* incluyen el deterioro del crecimiento postnatal, microcefalia congénita o postnatal, discapacidad intelectual grave con ausencia de desarrollo del habla, epilepsia, estereotipias y discinesia, patrones anormales de sueño, episodios inexplicables de llanto, reflujo gastroesofágico y aspiración recurrente (Kortum et al., 2011). Este fenotipo clínico definido también es denominado síndrome de *FOXG1* como equivalente a la designación original de variante congénita RTT. Sin embargo, el espectro fenotípico es complejo y todavía se está expandiendo (Vegas et al., 2018, Mitter et al., 2018).

El gen *FOXG1* se encuentra localizado en el cromosoma 14 (14q12) y está formado por un solo exón (Figura 6). Este gen codifica para una proteína que actúa como factor represor de la transcripción molecular con expresión restringida al tejido testicular y al cerebro fetal y del adulto (Ariani et al., 2008). En el período fetal, actúa sobre el neuroepitelio teleencefálico, el área nasal de la retina y el nervio óptico. Además, es fundamental para el desarrollo cerebral ya que en las regiones frontales favorece la proliferación de ciertos precursores neuronales e inhibe la neurogénesis prematura (Bahi-Buisson et al., 2010).

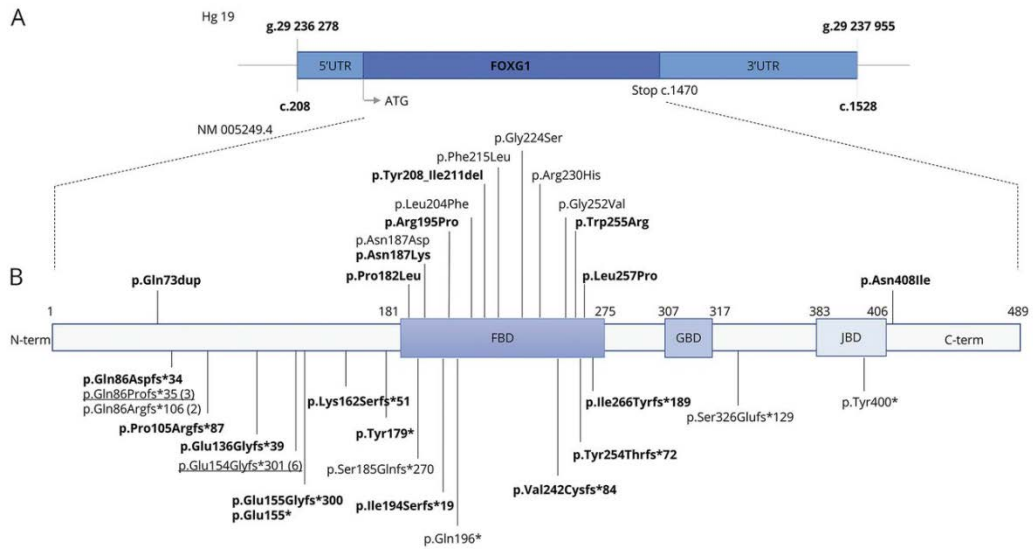


Figura 6. Estructura y distribución de las mutaciones en el gen *FOXG1*. (A) Representación esquemática del gen *FOXG1* y (B) estructura del dominio de la proteína FOXG1 y posiciones de las variaciones identificadas: N.term; dominio N-terminal; FBD, dominio de unión al ADN *forkhead*, 181–275aa; GBD, dominio de unión a Groucho, 307–317aa; JBD, dominio de unión a JARID1B, 383–406aa; C-Term, dominio C-terminal. *Imagen obtenida de Vegas et al., 2018*.

FOXG1 codifica para un factor de transcripción que contiene un dominio altamente conservado que abarca desde el dominio de unión *forkhead* (FBD) hasta el terminal C y una región N-terminal no conservada. Las mutaciones *FOXG1* incluyen variantes *frameshift*, deleciones y mutaciones puntuales (Vegas et al., 2018). Un estudio reciente sugiere que los fenotipos más severos se asocian con variantes que truncan la proteína en el extremo N-terminal y el FBD, y fenotipos más leves con variantes *missense* en el FBD. Las diferencias más significativas se relacionaron con el desarrollo motor y del habla, mientras que solo se encontraron diferencias límite con respecto a las anomalías del cuerpo calloso, la mielinización tardía y la microcefalia (Mitter et al., 2018).

2.3 Pacientes sin diagnóstico genético

En la actualidad hay más de 800 mutaciones diferentes descritas en el gen *MECP2* en más del 95% de los pacientes con RTT clásica y el 75% de los pacientes con RTT atípica (RettBASE: *MECP2* Variation Database: http://mecp2.chw.edu.au/). También las variantes atípicas del RTT como la variante congénita y la de epilepsia de inicio precoz se les han

asociado mutaciones en el gen *FOXG1 y CDKL5*, respectivamente (Kalscheuer et al., 2003, Ariani et al., 2008). Sin embargo, la etiología de un subconjunto de pacientes con diagnóstico clínico de RTT aún se desconoce.

Con el nacimiento de la secuenciación masiva (NGS; *next generation sequencing*), la etiología de estas pacientes y de muchos otros afectos con discapacidad intelectual (DI) severa ha podido ser resuelta. En estudios iniciales usando esta estrategia con grandes cohortes bien estudiadas de individuos afectos con discapacidad intelectual severa sin un diagnóstico previo positivo ya se podía apreciar un gran rendimiento diagnóstico del método. Además, esos resultados sugirieron que las SNV y CNV *de novo* que afectan a la región de codificación son una causa importante de DI severa (Gilissen et al., 2014). Centrándonos en la etiología de las pacientes RTT, diferentes grupos empezaron a realizar estudios de NGS de los casos negativos para el estudio de los genes que por el momento habían sido asociados al RTT (*MECP2, CDKL5* y *FOXG1*). El rendimiento diagnóstico de la técnica se encuentra entre el 50 y el 68% de positivos, donde una gran variedad de genes han sido asociados al RTT (Olson et al., 2015, Lopes et al., 2016, Srivastava et al., 2018, Iwama et al., 2019). Estos estudios han permitido mejorar la comprensión de las vías afectadas en los pacientes con RTT, y que en la gran mayoría de los genes alterados en estos pacientes se encuentran involucrados en las rutas de modulación de la cromatina, como las enzimas de modulación de la cromatina o en la función sináptica, necesaria para las sinapsis GABAérgica, glutamatérgica y dopaminérgica, el tráfico de vesículas sinápticas, la homeostasis iónica en las neuronas y el arrastre circadiano.

El vínculo de todos estos genes con las mismas vías podría explicar por qué los fenotipos de estos pacientes se superponen, causando una función sináptica deteriorada, trastornos del sueño y una desregulación importante de la expresión génica. Aún así queda mucho trabajo por hacer, ya que sigue habiendo pacientes si un claro diagnóstico genético que pueda ayudar a comprender mejor el desarrollo clínico que presentan y poder abordarlo con los tratamientos y terapias más adecuadas para su condición.

3. Evolución del diagnóstico genético

Debido a los rápidos avances en las tecnologías genómicas, los análisis genéticos se han vuelto esenciales en la práctica clínica y la investigación. Además, los métodos para el diagnóstico genético se han vuelto ampliamente accesibles y posibles de realizar incluso para laboratorios de menor tamaño y presupuesto. Durante estos años, el Instituto Nacional de Investigación del Genoma Humano (*National Human Genome Research Institute*; NHGRI) ha seguido y contabilizado los costes asociados a la secuenciación del ADN realizados en los centros de secuenciación financiados por el Instituto. Esta información ha servido como un punto de referencia importante para evaluar las mejoras en las tecnologías de secuenciación masiva del ADN (Figura 7). Hoy en día, con el uso cada vez más extendido de las técnicas de NGS, la mejora año a año de dicha tecnología y de las herramientas de análisis bioinformático, más pacientes pueden obtener un diagnóstico genético. Factor importante para un asesoramiento genético adecuado, la perspectiva futura del paciente y las posibles opciones de tratamiento que se les pueda aplicar.

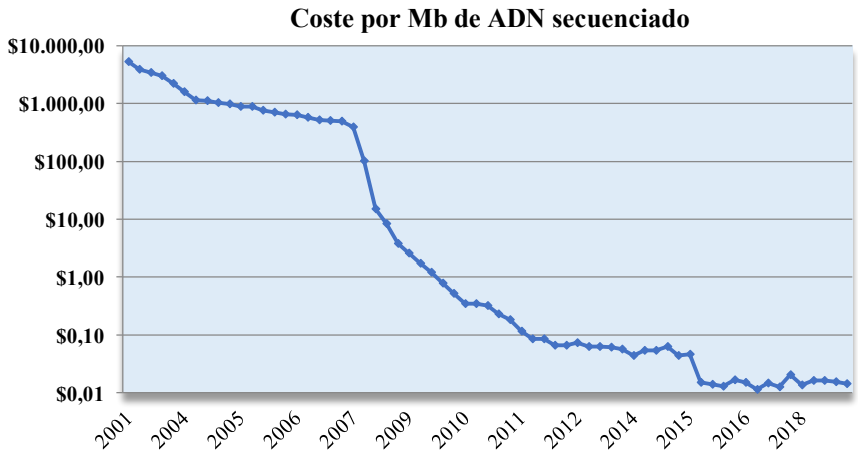


Figura 7. Datos de contabilidad de costos presentados por el NHGRI. Coste por megabase (Mb; un millón de bases) de secuenciación de ADN desde el año 2001 hasta la actualidad (Datos obtenidos de http://www.genome.gov/sequencingcosts).

El concepto NGS sirve para englobar todas aquellas tecnologías destinadas a llevar a cabo la secuenciación a gran escala de cualquier ácido nucleico (ARN y ADN), y consiste en secuenciar en paralelo y de forma simultánea millones de fragmentos de ADN, por lo que estos métodos de secuenciación también reciben el nombre de *High-throughput DNA sequencing*.

3.1 Fundamentos químicos de la NGS

El concepto básico detrás de la NGS es similar a la secuenciación Sanger. La ADN polimerasa cataliza la incorporación de desoxirribonucleótidos trifosfatos (dNTP) marcados con fluorescencia en una cadena molde de ADN durante ciclos secuenciales de síntesis de ADN. La diferencia crítica es que, en lugar de secuenciar un solo fragmento de ADN, este método extiende este proceso a través de millones de fragmentos en paralelo. Sin embargo, la preparación de muestras para su secuenciación por NGS no es tan sencilla como para la secuenciación por el método de Sanger.

El método de secuenciación por NGS requiere la preparación previa del ADN o ARN que se quiere analizar con la creación de las denominadas "librerías". Estas librerías consisten en fragmentos de ADN uniformes y con una longitud óptima para la plataforma de secuenciación que se vaya a utilizar. Para obtenerlas, existe una amplia variedad de protocolos de preparación:

1. Fragmentación: para la obtención los de fragmentos de ADN tanto la fragmentación mecánica (cizallamiento) como los métodos enzimáticos son adecuados para la preparación de librerías. Los métodos mecánicos permiten el cizallamiento aleatorio para producir una variedad de fragmentos superpuestos para cualquier región determinada del genoma, este método es ideal para ensamblar genomas sin una secuencia de referencia. Por otro lado, los métodos enzimáticos son relativamente rápidos y requieren menos inversión por adelantado, pero tienen algo de sesgo, ya que escinden algunas regiones preferentemente.

2. Ligación de adaptadores: una vez obtenidos los fragmentos, todos los métodos tienen en común la ligación de los adaptadores. Estos adaptadores son secuencias muy cortas y conocidas que sirven como sitios de unión de oligos universales durante los pasos de amplificación y secuenciación. Estos adaptadores, también sirven como "etiquetas" que nos permite saber de qué muestra es cada fragmento de ADN secuenciado, de este modo se pueden mezclar en un mismo ensayo el genoma de varios individuos para secuenciarlos simultáneamente en el secuenciador.

3. Selección por tamaño: la mayoría de protocolos usan perlas, denominadas *beads,* de inmovilización reversible en fase sólida, que proporcionan un método rápido y eficiente para la selección de los fragmentos adecuados. Sin embargo, la extracción de banda en gel todavía se usa comúnmente, ya que permite una selección de tamaño más precisa.

4. Amplificación: En la secuenciación del genoma completo, todos los fragmentos se secuencian, mientras que en la secuenciación del exoma completo o de regiones de interés solo se secuencia un subconjunto de todos fragmentos del genoma. En este subconjunto de fragmentos se lleva a cabo un enriquecimiento de la muestra, que consiste en amplificar y. captura, eliminando el resto de los fragmentos no deseados. Una vez finalizado este punto, lo que obtenemos son las llamadas librerías que se cargarán en el secuenciador masivo para llevar a cabo la secuenciación.

Las plataformas NGS varían considerablemente, pero son comunes en todos los ciclos de lavado y escaneo múltiples: se agregan nucleótidos, se genera una señal detectable tras la incorporación de nucleótidos a una cadena en crecimiento y los nucleótidos no incorporados se eliminan. Existen tres plataformas principales:

1. Roche: 454 Life Sciences Corp desarrolló la primera tecnología NGS y cambió fundamentalmente las percepciones de lo que podría lograrse con la secuenciación. Este método se basa en la pirosecuenciación, una técnica que se realiza mediante la adición cíclica de un nucleótido individual, ATP sulfurilasa y luciferasa. El pirofosfato se convierte en ATP mediante la ATP sulfurilasa, que actúa como sustrato para que la luciferasa genere una señal luminosa. Esta señal es proporcional al número de moléculas de pirofosfato liberadas y al número de nucleótidos incorporados. Cada señal de luz se procesa y se convierte en una secuencia de ADN mediante el *software* del instrumento.

2. Illumina: La tecnología Illumina fue desarrollada originalmente por la compañía Solexa y es una de las más usadas en la actualidad. En la plataforma de Illumina justo antes de los ciclos de secuenciación, los fragmentos se separan espacialmente y se unen a una superficie de cristal donde ocurre la secuenciación química. En esta superficie se encuentran cadenas de oligonucleótidos complementarios a las secuencias de los adaptadores que tienen los fragmentos de ADN de las librerías, por lo que se unirán por toda la superficie para su secuenciación. Seguidamente, se añaden cuatro nucleótidos marcados con fluorescencia al portaobjetos y compiten por incorporarse a las cadenas en crecimiento. En cada ciclo, los grupos se excitan con láser y la fluorescencia emitida se registra mediante un dispositivo de captura de imágenes. A medida que la posición de cada grupo individual permanece fija, el secuenciador crea un lapso de tiempo con las imágenes grabadas de todos los ciclos, y cada grupo genera una lectura.

3. Life Technologies (Ion Torrent): este método utiliza un enfoque muy similar la pirosecuencación 454 original. La secuenciación se realiza en los pocillos de un chip semiconductor en el que se pueden cargar perlas de PCR emulsionadas individuales. La secuenciación se realiza de la misma manera cíclica pero no hay enzimas adicionales y se utilizan nucleótidos naturales, en lugar de modificados con fluorescencia. A medida que se incorpora cada nucleótido, se liberan iones de hidrógeno, lo que cambia el pH de la solución en el pocillo. El chip detecta el cambio en el pH por un sensor de iones que se localiza en la parte inferior de cada pozo que lee los datos en un formato de diagrama de flujo similar al "pirograma" de Roche/454.

3.2 Fundamentos bioinformáticos de la NGS

Una vez finalizada la secuenciación se obtienen varios miles de fragmentos individuales sin ninguna información con respecto a la posición original de cada fragmento en el genoma, que se denominan los datos crudos, datos sin procesar que están almacenados en grandes archivos de texto. Estos archivos pueden pesar de decenas a cientos de gigabytes, conteniendo varios millones de secuencias de nucleótidos cortas denominadas lecturas que tienen una longitud de entre 35 a 400pb aproximadamente. Estos archivos crípticos deben someterse a un procesamiento computacional complejo para poder obtener datos significativos para el análisis, ya que el manejo manual de estos datos es bastante difícil.

El papel de la bioinformática es cada vez mayor en el manejo y análisis de esta enorme cantidad de datos biológicos que generamos a través de la investigación médica, biotecnológica y clínica a nivel mundial. Por lo que su implementación y mejora es muy importante para poder analizar la gran cantidad de datos obtenidos de manera eficiente y rápida para reducir costes y, sobretodo, tiempo de procesamiento.

3.2.1 Procesamiento de los datos crudos del secuenciador

Para determinar la posición de las lecturas en el genoma, deben alinearse (mapearse) a su ubicación más probable en el genoma humano de referencia y deben tenerse en cuenta las posibles discrepancias o lagunas. La alineación se basa únicamente en su secuencia, una tarea compleja cuando se trata de lecturas cortas de un genoma gigantesco. Idealmente, las lecturas deberían superponerse para cubrir cada base varias veces. Después de la etapa de alineación,

cada nucleótido se compara con su contraparte en el genoma de referencia y se registra, en un proceso conocido como llamada de variantes. Las diferencias con la referencia (desajustes, inserciones o huecos) se consideran mutaciones. En cualquier posición específica, se esperaría que un cambio homocigoto difiera del genoma de referencia en casi su totalidad de las lecturas, mientras que un cambio en heterocigosis estaría presente en el ±50% de las lecturas. La secuenciación y el mapeo no son procesos libres de errores, distinguir las variantes reales del ruido de fondo puede ser un desafío, por lo tanto, una gran profundidad de lectura que se denomina cobertura (número de lecturas diferentes que cubren una base específica en el genoma) es esencial para una llamada de variantes precisa y real.

3.2.2 Interpretación de los resultados obtenidos

El mayor desafío al que se enfrentan los genetistas al aplicar la NGS en el diagnóstico de enfermedades raras es determinar qué variantes raras, de las docenas o cientos detectadas, están potencialmente implicadas en el fenotipo del paciente. Por lo tanto, la priorización de variantes es un paso esencial, en el que se ha de ir con cautela en el proceso de diagnóstico genético de enfermedades raras.

Además de realizar el análisis *in silico* para predecir la patogenicidad de la variante, basados en la conservación de los nucleótidos/aminoácidos a lo largo de la cadena evolutiva y las diferencias entre las características fisicoquímicas del cambio de aminoácido como su polaridad y su carga, deben tenerse en cuenta tres conceptos esenciales para determinar la posible patogenicidad de un cambio (Roca et al., 2018).

1. La *tolerancia a la mutación* de los genes en los que se encuentran las variantes. Ciertos genes son mucho más tolerantes al fenómeno de la mutación que otros, en el desarrollo del proyecto de los 1000 Genomas se desveló que el grado de conservación varía entre genes (Genomes Project et al., 2012). Asimismo, varios estudios sobre la variación de la tasa de sustitución han demostrado que la tasa de mutación local varía según el genoma de los mamíferos. La tasa de sustitución se ve afectada por la proporción de GC (dado que la composición base no está equilibrada), el sitio en sí (CpG o no CpG) y el tipo de mutación (p. Ej., Transición, transversión, variantes de longitud) (Nachman and Crowell, 2000, Smith et al., 2002). Como se mencionó anteriormente, se han desarrollado muchas herramientas para tratar de priorizar variantes en términos de medidas de conservación a nivel filogenético o en función

de las características de los aminoácidos. Sin embargo, se han desarrollado pocos enfoques análogos para priorizar los genes en los que se detectan las variantes (Feng, 2017, Kennedy et al., 2014). Petrovski et al. (2013) demostraron que los genes responsables de las enfermedades mendelianas son significativamente menos tolerantes a la variación genética funcional que los genes que no causan ninguna enfermedad conocida, y observaron una sorprendente variación en la intolerancia entre los genes que causan diferentes clases de enfermedades genéticas (Petrovski et al., 2013). Por lo tanto, una puntuación a nivel de gen podría combinarse con las puntuaciones a nivel de variante bien establecidas para resaltar mutaciones causales candidatas.

2. La *arquitectura mutacional* de cada gen. Consiste en evaluar, para una puntuación de tolerancia dado, los tipos de mutaciones y su posición dentro del gen, y la asociación entre estas variantes y diferentes fenotipos. Ha sido descrito con fiabilidad que la perturbación de diferentes regiones funcionales de un gen puede afectar a diferentes funciones y/o procesos, dando lugar a fenotipos diversos (Gussow et al., 2016). Además, los diferentes tipos de mutación (las que truncan la proteína frente a las *missense* o las de ganancia de función frente a las de pérdida) pueden dar lugar a fenotipos distintos. El conocimiento de la arquitectura mutacional específica de cada gen nos permite priorizar las variantes detectadas por el análisis NGS. La evaluación de la posible participación de una variante en un fenotipo dado debe considerar la arquitectura mutacional específica del gen en el que se encuentra la mutación. La misma variante (tipo y ubicación) podría tener efectos muy perjudiciales en un gen, pero no en otro. Ciertos genes son susceptibles solo a variantes que truncan la proteína, pero toleran variantes *missense* (por ejemplo, *TTN* y *SYNE1*). De manera similar, algunos genes como *KCNQ2* solo son susceptibles a mutaciones *missense* que crean una ganancia de función en la proteína, mientras que las mutaciones que truncan la proteína dando una pérdida de función dan lugar a fenotipos mucho más leves (Kato et al., 2013, Weckhuysen et al., 2013). Asimismo, en el caso de *CDKL5*, como se ha mencionado anteriormente, presenta ciertos exones específicos del gen sensibles a mutaciones. Tener información actualizada sobre la arquitectura mutacional específica de cada gen permite priorizar correctamente las variantes detectadas por el análisis NGS.

3. El *modelo de herencia* de las variantes detectadas. El conocimiento del modo de herencia de las variantes es importante al evaluar su participación en el fenotipo clínico de un paciente. En casos de enfermedades graves dominantes, es esencial determinar si la variante detectada es una mutación heredada o *de novo*. Del mismo modo, es esencial determinar si las variantes detectadas en un gen recesivo se encuentran en alelos opuestos antes de asociarlos con la enfermedad. Los estudios de WGS basados en familia han demostrado que, en promedio, 74 SNV de línea germinal ocurren *de novo* en el genoma de un individuo y una mutación *de novo* por exoma (Conrad et al., 2011). La combinación de WES con un diseño trío paciente-progenitores permite la detección de un promedio de 1.68 mutaciones *de novo* por paciente, lo que a menudo permite una identificación rápida del gen subyacente a la condición del paciente (Rauch et al., 2012). Las mutaciones *de novo* son la forma más extrema de variación genética rara. En general, son más perjudiciales que las variaciones heredadas ya que han sido sometidos a una selección evolutiva menos estricta (Eyre-Walker and Keightley, 2007). Los genes dominantes intolerantes son aquellos que tienen más probabilidades de verse afectados por mutaciones *de novo*. Cualquier alteración en un gen intolerante a la mutación resulta en un daño tal que no puede perpetuarse a nivel de la población, el individuo no puede transmitirlo a su descendencia porque sus efectos les impiden reproducirse. Por lo tanto, en los genes dominantes más conservados, solo las mutaciones *de novo* están asociadas con patologías. De hecho, en los casos más extremos, solo se permiten mutaciones de mosaico *de novo*. El papel destacado de las mutaciones *de novo* en enfermedades raras se ha revelado en los últimos años gracias al uso de WES en tríos (Veltman and Brunner, 2012).

Por todo ello, un grupo de trabajo compuesto por representantes del Colegio Americano de Genética Médica (ACMG), la Asociación de Patología Molecular (AMP) y el Colegio de Patólogos Americanos, ha desarrollado unos estándares y guías para una correcta priorización y clasificación de variantes detectadas en el genoma. Estos estándares y pautas de la ACMG se desarrollaron principalmente como un recurso educativo para los genetistas de laboratorio clínico para ayudar a proporcionar servicios de laboratorio clínico de calidad (Richards et al., 2015). Estas recomendaciones se aplican principalmente a la variedad de pruebas genéticas utilizadas en laboratorios clínicos, incluidos genotipados, genes individuales, paneles, exomas y genomas. Este informe recomienda el uso de terminología estándar específica: "patogénica", "probablemente patogénica", "significado incierto",

"probablemente benigna" y "benigna", para describir las variantes identificadas en los genes que causan trastornos mendelianos. Además, estas recomendaciones describen un proceso para clasificar variantes en estas cinco categorías en función de los criterios y de evidencias que presentan las variantes detectadas (p Ej., datos de frecuencias poblacionales, computacionales, funcionales, de segregación) (Figura 8).

	Benigno		Patogénico			
	Fuerte	**Soporte**	**Soporte**	**Moderado**	**Fuerte**	**Muy fuerte**
Datos Poblacionales	La MAF es demasiado alta para el trastorno O la detección en controles no es consistente con la penetrancia de la enfermedad			No presente en las bases de datos poblacionales	Prevalencia en individuos afectos mayor estadísticamente que en controles	
Datos predictivos y computacionales		Múltiples líneas computacionales sugieren que no hay impacto en el gen Missense en genes que solo las variantes que truncan causan enfermedad Cambios sinónimos que no predicen un efecto en el splicing Indels in-frame en zona repetitiva sin función conocida	Múltiples líneas computacionales sugieren que hay un efecto deletéro en el gen	Missense noveles en una posición donde otra missense ha sido descrita como patogénica Variante que modifique la longitud de la proteína	Mismo cambio aminoacídico ya establecido como patogénico	Variante que trunque en un gen donde la pérdida de función es un mecanismo conocido para desarollar enfermedad
Datos Funcionales	Estudios funcionales bien establecidos no demuestran ningún efecto deletéreo en el gen		Missense en gen con bajo número de variantes missense benignas y comunes las patogénicas	*Hot spot* para mutaciones o zona con dominio funcional bien estudiado sin variantes benignas	Los estudios funcionales bien establecidos muestran un efecto deletéreo	
Segregación	No cosegrega con la enfermedad		Cosegregación en múltiples individuos afectos en una familia →	Datos de segregación incrementados →		
de novo				*de novo* (sin confirmación en los progenitores)	*de novo* (con confirmación de los progenitores)	
Alelos		Observado en *trans* con una variante dominante Observado en *cis* con una variante patogénica		Para trastornos recesivos, cambio detectado en *trans* con una variante patogénica		
Otras bases de datos		Fuente de buena reputación lo considera benigno	Fuente de buena reputación lo considera patogénico			
Otros		Encontrado junto a otra variante que puede ser la causa	Fenotipo del paciente claramente asociado con el gen			

Figura 8. Marco de evidencias. Este cuadro organiza cada uno de los criterios según el tipo de evidencia, así como la fuerza de los criterios para una afirmación benigna (lado izquierdo) o patogénica (lado derecho). MAF: *Minor allel frequency*, *cis*: 2 cambios sobre el mismo alelo del gen; *trans*: dos cambios sobre alelos distintos del gen. *Figura modificada de Richards et al., 2015.*

3.3 Aproximación de la metodología NGS al diagnóstico del RTT

Con la posibilidad de *multiplexar* genes y pacientes, y secuenciarlos al mismo tiempo, la rentabilidad de la técnica es comparable al análisis de secuenciación de Sanger de un único gen (Koboldt et al., 2013). Sin embargo, aún no se ha establecido una estrategia de implementación óptima para los laboratorios de diagnóstico asistencial (Di Resta et al., 2018). Básicamente, existen cuatro enfoques en la NGS para la secuenciación del ADN para mejorar el rendimiento diagnóstico en grupos de enfermedades enormemente heterogéneos:

- Enriquecimiento dirigido a uno o muy pocos genes. Con la secuenciación masiva de un subconjunto de fragmentos muy limitado del genoma, llegamos a conseguir coberturas del 100% de estos genes con profundidades de lectura de miles, pudiendo detectar de este modo mosaicismos. Esta aproximación se utiliza cuando las características clínicas y otros resultados de pruebas del paciente son consistentes para un trastorno particular en el que se puede establecer una asociación clara con un gen específico. También hay una clara ventaja interpretativa en el enfoque de un solo gen porque la probabilidad de descubrir múltiples variantes de significado clínico incierto (VUS) es mínima. Por ejemplo, el gen *FGFR3* es el único que se conoce que está asociado con la acondroplasia. La prueba de un solo gen de *FGFR3* detecta mutaciones en el 99% de los pacientes con acondroplasia y, por lo tanto, es el enfoque más eficiente en términos de coste-tiempo (Richard M Pauli, and Janet M Legare. Achondroplasia *GeneReviews*. [On line]). Sin embargo, el médico debe tener conocimiento de los hallazgos diagnósticos clínicos y radiográficos de la acondroplasia para seleccionar la prueba genética correcta.

- Enriquecimiento dirigido de un subconjunto de genes (paneles dirigidos). Se centran en la secuenciación de regiones específicas de interés o un subconjunto de genes con asociaciones conocidas o sospechas con la enfermedad o fenotipo de estudio. Este enfoque suele ser la primera línea de pruebas, para el RTT tanto clásico como atípico. Además, los paneles se pueden personalizar y optimizar para diferentes regiones y tipos de muestra, lo que permite la determinación de variantes de nucleótidos individuales (SNV) a partir de NGS de una manera más rentable. Los paneles específicos están mejorando constantemente porque con la investigación básica, exoma y genoma de pacientes sin diagnóstico genético, se

descubren nuevos genes y sus funciones y, posteriormente, se asocian con enfermedades humanas. Por esta razón, los paneles específicos son el mejor enfoque en términos de diagnóstico genético.

- Secuenciación del exoma completo (WES). Este tipo de prueba a menudo implican el estudio de el paciente y a ambos progenitores (trío) para ayudar en la interpretación de variantes. El desafío actual de WES es discernir entre variantes benignas y patogénicas, ya que el exoma completo de una persona sana revela aproximadamente 30.000-100.000 variaciones si se compara con el genoma de referencia estándar. Utilizando bases de datos y herramientas de *software* permiten clasificar las variantes según su nivel de patogenicidad. Sin embargo, este método es incierto y los estudios funcionales siguen siendo necesarios para demostrar completamente la patogenicidad de las variaciones encontradas. Asimismo, las mutaciones en las regiones intrónicas y promotoras no están cubiertas y tampoco se pueden detectar variantes estructurales como inversiones y translocaciones.

- Secuenciación del genoma completo (WGS). Actualmente, para el estudio de enfermedades mendelianas como el RTT, la secuenciación del genoma completo no se utiliza como primera aproximación debido a la complejidad de la construcción de la librería y su secuencación y la dificultad e incertidumbre del análisis de datos, ninguno de los cuales "encaja" en el paradigma de laboratorio clínico. Lo más probable es que, en los casos en que se requiera la secuenciación y el análisis del genoma completo, que pueden ser aquellos en los que el análisis del exoma completo haya dado un resultado negativo, la producción y el análisis de datos se realizarán en un laboratorio de investigación con la experiencia necesaria, y cualquier hallazgo clínicamente relevante será verificado posteriormente por un laboratorio clínico aprobado utilizando métodos convencionales y por sus correspondientes estudios funcionales.

Por lo tanto, la implementación global de esta tecnología en el laboratorio de investigación ha llevado a un aumento importante en la identificación de nuevos genes relacionados con el fenotipo RTT / RTT-*like* (Schonewolf-Greulich et al., 2019, Wang et al., 2018).

4. Tratamientos y nuevas terapias génicas

Tras la identificación de *MECP2* como el principal factor genético subyacente a RTT (Amir et al., 1999), se ha centrado un enorme esfuerzo para intentar moldear la enfermedad utilizando modelos animales como el ratón y los modelos derivados de células iPSC.

4.1 Modelos animales del RTT

Mecp2 se encuentra en todos los vertebrados, pero no en organismos modelo no vertebrados, incluida la mosca de la fruta o el gusano (Hendrich and Tweedie, 2003). Por lo tanto, el desarrollo de modelos de ratón era necesario para una comprensión mecanicista del inicio y la gravedad de los síntomas clínicos. Poco después de la identificación de *MECP2* como el gen causante del RTT, se generaron dos modelos de ratones nulos para Mecp2, que ahora son los principales modelos utilizados para estudiar la enfermedad. La línea de ratones $Mecp2^{tm1.1Bird}$ carecen completamente del producto proteico (Guy et al., 2001), mientras que la línea $Mecp2^{tm1.1Jae}$ expresa la proteína truncada (Chen et al., 2001). Sin embargo, ambos modelos nulos muestran un fenotipo similar que recapitula los síntomas del RTT y ambos se han utilizado ampliamente para estudiar la fisiopatología de la enfermedad. En la actualidad hay un gran número de modelos murinos para estudio del RTT (Vashi and Justice, 2019).

Los ratones machos nulos para Mecp2 ($Mecp2^{-/y}$) han sido usados con mayor frecuencia que los ratones hembras heterocigotos ($Mecp2^{+/-}$), ya que los primeros muestran síntomas graves de inicio temprano similares a los de los pacientes con RTT, como irregularidades respiratorias y anormalidades motoras, junto con una letalidad temprana (Guy et al., 2001, Chen et al., 2001). Los ratones hembra $Mecp2^{+/-}$ proporcionan una representación genética más precisa del trastorno, aunque su sintomatología neurológica se desarrolla más tarde y muestran una mayor variabilidad fenotípica en comparación con los ratones macho $Mecp2^{-/y}$ (Samaco et al., 2013). Sin embargo, deben usarse los ratones hembras, así como los ratones macho, en estudios preclínicos para potenciales terapéuticos ya que son mujeres las que mayoritariamente presentan la enfermedad (Ure et al., 2016).

Los modelos de ratones mutantes Mecp2 recapitulan el amplio espectro de fenotipos observados en pacientes con RTT (Figura 7). Los síntomas más profundos tanto en pacientes como en los modelos de ratón son las anomalías motoras, que incluyen movilidad reducida, coordinación motora deteriorada, marcha atáxica y temblores. Mientras que los pacientes

humanos reemplazan el uso intencional de las manos con movimientos estereotípicos, los modelos de ratones se agarran de las extremidades posteriores (Guy et al., 2001, Chen et al., 2001). Morfológicamente, tanto los pacientes humanos como los modelos de ratones Mecp2 exhiben un volumen cerebral reducido e hipotrofia neuronal (Chahrour and Zoghbi, 2007). Conductualmente, los pacientes con RTT experimentan regresión neurológica y pérdida del habla, dos rasgos que no pueden medirse en modelos de ratón. Sin embargo, ambos pueden experimentar algunos déficits de aprendizaje (Elefant and Wigram, 2005). Curiosamente, mientras que los pacientes con RTT exhiben típicamente mayor ansiedad y evitación social (Mount et al., 2002), los ratones son menos ansiosos y más sociales, según lo evaluado por las pruebas fenotípicas disponibles para roedores (Samaco et al., 2013, Orefice et al., 2016). Finalmente, ambos experimentan irregularidades respiratorias (Ramirez et al., 2013), anomalías cardíacas (intervalos QTc prolongados) (McCauley et al., 2011, Hara et al., 2015), convulsiones y una esperanza de vida más corta (Guy et al., 2001). En general, los ratones mutantes Mecp2 exhiben una amplia gama de fenotipos que recapitulan los síntomas de pacientes humanos con RTT, lo que los convierte en excelentes modelos para estudiar el trastorno (Vashi and Justice, 2019).

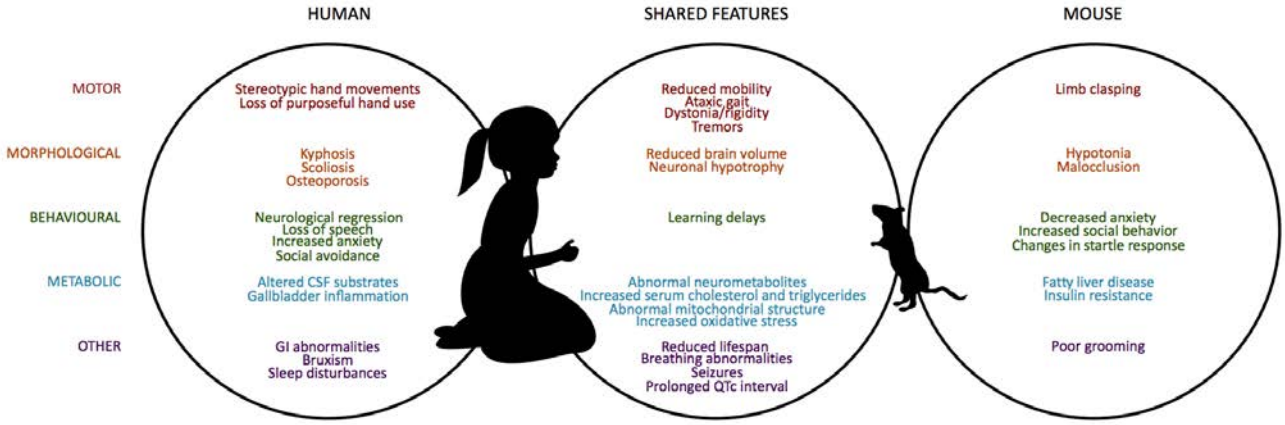


Figura 9. Características compartidas entre pacientes con RTT y modelos de ratón Mecp2. *Imagen obtenida de Vashi and Justice, 2019.*

Los modelos de ratón para mutaciones *missense* (p.T158A, p.T158M y p.R306C) recapitulan fenotipos similares al RTT. Las variantes p.T158A y p.T158M que afectan el dominio MBD, reducen la estabilidad de la proteína y causan reducciones en la amplitud de los potenciales de respuesta a eventos en la corteza auditiva en los modelos de ratón (Cuddapah et al., 2014, Brown et al., 2016). La variante p.R306C, que se localiza en el dominio TRD de la proteína, evita que Mecp2 interactúe con el complejo correpresor NCoR-

SMRT (mediador silenciador del ácido retinoico y el receptor de la hormona tiroidea) e inhibe la represión transcripcional (Lyst et al., 2013, Brown et al., 2016). Estas tres mutaciones son las variantes *missense* patogénicas más frecuentes descritas en *MECP2* por lo que estos modelos son muy relevantes para el estudio de la fisiopatología del RTT y son herramientas valiosas para estudios preclínicos.

Una limitación de los modelos de ratones transgénicos es que la aparición de síntomas neurológicos ocurre mucho más tarde que en el desarrollo que en los humanos. Además, los fenotipos relacionados con RTT pueden ser menos severos en roedores. Se deben considerar otros factores, incluidos los antecedentes genéticos y las condiciones de cría, al analizar los fenotipos expresados. Además, se han desarrollado modelos RTT en otras especies, incluidas ratas y primates no humanos. Estos recapitulan las características clave de RTT y ofrecen características de comportamiento únicas, incluida la regresión de una habilidad psicomotora aprendida en ratas (Veeraragavan et al., 2016) y medidas únicas de seguimiento ocular en primates (Chen et al., 2017).

4.2 Estrategias terapéuticas

Con los grandes avances realizados hasta la fecha para detectar y reconocer componentes que se encuentran en las vías de señalización de *MECP2* se han identificado ciertas dianas con un gran potencial terapéutico. Estas dianas pueden servir para mejorar la señalización de los neurotransmisores, la señalización del factor de crecimiento y el metabolismo. Es importante destacar que varios de estos tratamientos ya han sido realizados en modelos de ratón cuyos síntomas mejoraron y ya han generado ensayos clínicos en pacientes (Vashi and Justice, 2019).

Sin embargo, estas estrategias también han demostrado presentar varias limitaciones. Primero, debido a que aún se desconocen las funciones precisas y bien acotadas de *MECP2*, es difícil identificar todas las vías que regula, especialmente aquellas que pueden cambiar sutilmente en ausencia de *MECP2*. Otro desafío es la especificidad: *MECP2* se recluta en señales de metilación del ADN, que son altamente específicas para cada tipo celular (Deaton et al., 2011). Por lo tanto, los genes reguladores de MeCP2 probablemente variarán según el estado de metilación y en los diferentes tipos celulares. Consistentemente, un estudio reciente encontró que la regulación errónea de los genes dentro de los subtipos de neuronas en

ratones hembra mutantes Mecp2$^{+/-}$ depende en gran medida de las marcas epigenéticas específicas de cada tipo celular (Renthal et al., 2018). Finalmente, es difícil discernir qué vías se ven afectadas como el resultado primario de la pérdida de MeCP2 y cuáles son los efectos secundarios que se desarrollarán más adelante que probablemente tengan menos valor como objetivo terapéutico. Sin embargo, los avances en las técnicas de biología molecular están facilitando la resolución de estos problemas. Por ejemplo, la secuenciación de ARN de una sola célula puede identificar genes mal regulados en células individuales deficientes en MeCP2 y la inmunoprecipitación de cromatina de los sitios de unión conocidos de *MECP2* podría identificar sus objetivos transcripcionales directos en un tipo de célula específico. Vashi and Justice, 2019, destacan y resumen los tratamientos preclínicos dirigidos a vías posteriores a MeCP2 que han llevado a ensayos clínicos en pacientes con RTT (Annexo Tabla 1).

4.1.1 Señalización del factor de crecimiento como tratamiento para RTT

Estudios iniciales de las vías de señalización de *MECP2* identificaron el gen que codifica para el factor neurotrófico derivado del cerebro (*brain-derived neurotrophic factor*; *BDNF*) como una de sus dianas transcripcionales (Chen et al., 2003) y están siendo objeto de terapias (Katz et al., 2016). *BDNF* es un miembro de la familia de factores de crecimiento de la neurotrofina, que se une a la quinasa B relacionada con tropomiosina (TrkB) para estimular las cascadas de señalización involucradas en el crecimiento de neuritas, la función sináptica y la diferenciación neuronal (Amaral and Pozzo-Miller, 2007). Los niveles de BDNF se encuentran reducidos en ratones mutantes Mecp2 (Chang et al., 2006). Y, aunque la administración directa de BDNF no es un tratamiento factible ya que no es capaz de cruzar la barrera hematoencefálica, el fingolimod (FTY720), su análogo, es capaz de aumentar los niveles de BDNF y mejorar la actividad motora en ratones mutantes Mecp2 (Deogracias et al., 2012). Aunque hasta la fecha no ha habido ensayos clínicos con fingolimod en humanos.

Asimismo, el factor de crecimiento similar a la insulina 1 (*insulin-like growth factor 1*; IGF1) se encuentra también regulado transcripcionalmente por *MECP2*, pero este sí puede traspasar la barrera hematoencefálica (Itoh et al., 2007). Los ratones Mecp2$^{-/y}$ tratados con IGF-1 muestran una función locomotora mejorada, menos irregularidades respiratorias y una prolongación de la vida. IGF-1 también hace que aumente el peso cerebral de los ratones al mismo tiempo que restaura la densidad de la columna neural en la corteza motora y mejora

la transmisión sináptica excitadora en las neuronas de la corteza sensitivomotoras (Tropea et al., 2009). El tratamiento con IGF1 ha demostrado su eficacia en un ensayo en fase I/IIa (Khwaja et al., 2014). Sin embargo, los ensayos clínicos posteriores no produjeron mejorías significativas de los síntomas (O'Leary et al., 2018). También ha sido desarrollado un análogo sintético derivado de IGF-1, NNZ2566 (Trofinetide) que inhibe la neuroinflamación, restaura la función glial, corrige los déficits sinápticos y regula la respuesta al estrés oxidativo en ratones modelo para el síndrome del X frágil (Deacon et al., 2015). Asimismo, un ensayo clínico en fase II de NNZ-2566 en pacientes RTT ha demostrado una mejoría en la respiración y el comportamiento motor de las pacientes, que pronto avanzará a los ensayos de fase III (Ip et al., 2018).

4.1.2 Tratamientos potenciales dirigidos a la señalización de neurotransmisores

Ya que se ha observado un defecto en la señalización de neurotransmisores en modelos de ratón nulos para Mecp2, tratamientos dirigidos a estas vías han sido objeto de estudio en ratones, y en algunos casos han sido probados para ensayos clínicos.

La desipramina, un inhibidor de la recaptación de noradrenalina, mejoró las irregularidades respiratorias y las apneas en ratones Mecp2$^{-/y}$ (Roux et al. 2007; Zanella et al. 2008). Sin embargo, en un ensayo clínico posterior no se pudieron detectar mejorías clínicas en pacientes con RTT tratadas con este fármaco (Mancini et al. 2018). Sin embargo, la serotonina 1, el agonista y receptor similar a la dopamina D2, redujeron la apnea respiratoria en un 15-30% en ratones mutantes de Mecp2, pero no tuvieron ningún efecto sobre la actividad motora (Abdala et al. 2014). Como resultado, actualmente se está probando su efectividad para mejorar los síntomas respiratorios en RTT (NCT02790034). Finalmente, la ketamina, un agonista del receptor de NMDA, se probó en dos laboratorios diferentes por su potencial terapéutico en RTT (Kron et al. 2012; Patrizi et al. 2016). Estos estudios preclínicos mostraron que la dosis baja de ketamina podría aumentar la actividad en la red cortical al tiempo que disminuye la excitabilidad sináptica en la red del tronco encefálico, apuntando a un posible desequilibrio en la actividad neuronal en todo el cerebro en ratones mutantes para Mecp2. El tratamiento con ketamina mejoró la movilidad de las extremidades, la coordinación motora y las apneas respiratorias en ratones Mecp2. La seguridad de la ketamina se está evaluando actualmente en pacientes con RTT (NCT03633058).

4.1.3 Los defectos metabólicos en RTT pueden ser dirigidos terapéuticamente

Hay publicados estudios adicionales que revelaron que el metabolismo de los lípidos se encuentra gravemente desequilibrado en los modelos de ratones mutantes para Mecp2. Se ha descrito que antes del inicio de la sintomatología, el colesterol cerebral ya se encuentra notablemente elevado los modelos de ratón, y en suero e hígado el colesterol y los triglicéridos se encuentran elevados. Sorprendentemente, las estatinas hipolipemiantes, fármacos utilizados para reducir el colesterol y los triglicéridos, regulan los niveles de lípidos, mejoran los síntomas motores y prolongan la vida útil en ratones (Buchovecky et al. 2013). Además, también se han observado niveles elevados de lípidos en un subconjunto de pacientes con RTT, lo que indica que la reutilización de los medicamentos con estatinas puede ser una opción de tratamiento viable para beneficiar a las pacientes (Justice et al. 2013; Segatto et al. 2014). Las anomalías en la homeostasis del colesterol están asociadas con muchas enfermedades neurológicas (Tint et al. 1994; Puglielli et al. 2003; Bi y Liao 2010; Berry-Kravis et al. 2015), y por lo tanto el RTT no es una excepción. Es importante destacar que estos hallazgos llevaron a un ensayo clínico que probó la eficacia y seguridad de las estatinas en pacientes con RTT (NCT02563860). Los defectos metabólicos en RTT no se limitan a la síntesis de colesterol. Tanto los pacientes con RTT como los modelos de ratones mutantes para Mecp2 muestran una estructura y función mitocondrial anormales (Eeg-Olofsson et al. 1988; Dotti et al. 1993; Kriaucionis et al. 2006; Gold et al. 2014). Es importante destacar que la función mitocondrial comprometida puede afectar en gran medida a la producción de energía celular. Si bien las mitocondrias son necesarias en todas las células del cuerpo, su papel es especialmente importante en los tejidos con altas demandas de energía, como los nervios y los músculos. El RTT comparte muchas características de las enfermedades mitocondriales, incluyendo inicio sintomático temprano, retraso del desarrollo, regresión neurológica, tono muscular deficiente, convulsiones y problemas gastrointestinales (Schon y Manfredi 2003). Consistentemente, tanto los pacientes con RTT como los ratones mutantes de Mecp2 presentan un aumento del estrés oxidativo y una disminución de los niveles de enzimas mitocondriales (Haas et al. 1995; Kriaucionis et al. 2006; De Felice et al. 2009; Leoncini et al. 2011). En ratones mutantes Mecp2, los marcadores de estrés oxidativo aumentan con la edad, lo que sugiere una disfunción progresiva en la función mitocondrial (De Felice et al. 2014). Es importante destacar que la producción de energía mitocondrial está estrechamente relacionada con la síntesis del colesterol. Varios pasos en la biosíntesis de colesterol requieren fuentes de ATP mitocondrial como donante

de electrones en las reacciones de oxigenación. Por lo tanto, las perturbaciones en el colesterol pueden estar relacionadas con la disfunción mitocondrial en el RTT. Curiosamente, las sustancias anapleróticas pueden reponer compuestos intermedios en la vía de producción de energía, mejorando la producción de energía mitocondrial. La dieta complementada con la triheptanoína anaplerótica ha revertido los síntomas de algunos trastornos metabólicos al corregir la producción de energía (Roe et al. 2002; Mochel et al. 2005). Esta estrategia fue adaptada para su uso en modelos de ratón de RTT. Sorprendentemente, los ratones mutantes Mecp2 alimentados con una dieta con suplementos de triheptanoína mostraron una mejor morfología mitocondrial y una mejor producción de energía. Esto se tradujo en una mejor coordinación motora y una mayor vida útil en estos ratones. Dos ensayos clínicos que investigan la eficacia de la triheptanoína para mejorar las convulsiones, el tono muscular y la mejora de los síntomas en pacientes con RTT se encuentran actualmente en sus etapas iniciales (NCT02696044 y NCT03059160).

4.3 Terapias génicas

Los productos farmacéuticos no siempre pueden reparar un mal funcionamiento del cuerpo humano en su totalidad. A veces, la única forma de tratar por completo las afectaciones de un individuo es con la manipulación de su genoma: la base en la que se construyen los sistemas biológicos de cada ser. Algunos investigadores están utilizando técnicas de edición de genes como *CRISPR* para modificar con precisión las secuencias de la región genómica que interesan. Otros están modificando genéticamente las células inmunes para inculcarles la capacidad de combatir el cáncer. Y en los últimos años, ha habido una aceleración exponencial en el desarrollo de una amplia gama de tratamientos en los que los genes que no funcionan correctamente y son causantes de enfermedades se reemplazan en su totalidad por otros que cumplen su función biológica.

La meta y vía potencial final del tratamiento y la cura para el RTT implica la introducción de una copia normal de *MECP2* en las células mediante terapia génica. La terapia génica es un enfoque prometedor para el tratamiento de muchos trastornos y ha tenido éxito en revertir los síntomas en modelos de ratón de fibrosis quística, hemofilia, síndrome de Hunter, diabetes, obesidad, ELA y más. Recientemente, el primer tratamiento dirigido de terapia génica fue aprobado en América del Norte para tratar a pacientes con amaurosis congénita de Leber, una enfermedad ocular hereditaria rara (Kumaran et al. 2018).

Para tratar el RTT, los enfoques de terapia génica deben utilizar un vector apropiado capaz de cruzar la barrera hematoencefálica y poder transducir muchas células en el SNC, y además ser capaz de mantener una expresión estable a largo plazo del *MECP2* necesario. Asimismo, se deben desarrollar estrategias para evitar la represión transgénica y evitar su sobreexpresión. Al igual que la reactivación del Xi, la terapia génica se dirige a todas las células independientemente del estado de mutación *MECP2*. Por lo tanto, las células que expresan MeCP2 normal tendrán 2 veces el nivel normal de la proteína, lo que conducirá a posibles efectos tóxicos por su sobreexpresión. Una preocupación adicional de la terapia génica es que a menudo se necesitan altos niveles virales para infectar una gran proporción de células, pero las células que inadvertidamente reciben más de una partícula viral y, por lo tanto, más de una copia de *MECP2* se sobrecargarían. Las estrategias para eludir estos problemas podrían proporcionar un factor como un siARN (*small interfering RNA*) para suprimir la expresión del *MECP2* endógeno de modo que solo se exprese el *MECP2* transgénico (Gadalla et al. 2011). Sin embargo, esto requeriría que el *MECP2* transgénico posea un promotor sin fisuras para imitar completamente la expresión génica endógena. Además, miARNs (*microRNAs*) han sido utilizados para controlar la expresión transgénica exógena al mediar la degradación de los ARNm transgénicos de una manera específica en tejido y pueden ser beneficiosos en los enfoques de terapia génica en RTT para limitar la toxicidad fuera de su diana (Geisler y Fechner 2016).

Los vectores de virus adeno-asociados recombinantes (AAV) se han utilizado en estudios preclínicos de terapia génica por su capacidad para cruzar la barrera hematoencefálica, infectar neuronas y mediar la expresión estable a largo plazo del transgen sin inflamación ni toxicidad (Gonçalves 2005; Foust et al. 2009). El AAV tiene un genoma de ssADN (*single strand* DNA) de 4,7 kb del que se pueden eliminar 4,4 kb de ADN viral y reemplazarlo con un transgen humano. Los vectores AAV autocomplementarios tienen una eficiencia de transducción de 10 a 100 veces mayor, pero su inconveniente es que su capacidad de almacenar se reduce a la mitad, solamente 2.2 kb, lo que dificulta el empaquetado de genes grandes (McCarty et al. 2001). Fueron realizados ensayos donde se inyectó en ratones machos $Mecp2^{-/y}$ el vector AAV9-MECP2 bajo el promotor de la β-actina de pollo y se observó una eficiencia de transducción de las regiones cerebrales que varió del 7 al 42% de células infectadas, con la tasa de infección más alta en el hipotálamo y la más baja en el cuerpo estriado. Sin embargo, esta baja eficiencia de infección fue suficiente para aumentar la vida útil de los ratones a 16,6 semanas y mejorar las deficiencias motoras, pero no tuvo ningún

efecto sobre los síntomas respiratorios. Los ratones de tipo salvaje que recibieron inyecciones neonatales del mismo vector no mostraron déficits significativos, lo que sugiere una tolerancia para la sobreexpresión moderada de MeCP2. Por otro lado, para probar un enfoque de reemplazo de *MECP2* encarado a humanos se inyectó por vía intravenosa un vector de virus adenoasociado autocomplementario diseñado para impulsar la expresión de MeCP2 a partir de un fragmento del promotor Mecp2 (scAAV9-MECP2) en ratones juveniles Mecp2$^{-/y}$. Sin embargo, su eficiencia de transducción cerebral fue baja (~ 2-4% de las neuronas), todavía se observaron mejoras modestas en la supervivencia. Por lo tanto, los estudios futuros deberían aumentar la especificidad de la terapia AAV9 antes de trasladarse a la clínica. De todos modos, estos resultados apoyan el concepto de terapia génica en *MECP2* para pacientes RTT. (Gadalla et al. 2013).

Un estudio posterior en el cuál se empezaron a usar ratones hembra para realizar los ensayos, el modelo más apropiado dada la naturaleza de la enfermedad. En este estudio mostraron que scAAV9-MECP2 bajo el control de un fragmento de su propio promotor es capaz de estabilizar o revertir significativamente los síntomas cuando se administra sistémicamente en ratones RTT hembra (Garg et al. 2013). Más adelante se realizaron estudios para investigar el impacto y la eficacia del diseño del vector, la dosificación y la ruta de administración y determinar la seguridad de la terapia génica de aumento de dosis en modelos de ratón RTT. Sus resultados mostraron que su administración se asoció con una posible ventana terapéutica estrecha y creó toxicidad hepática en dosis altas. Por otro lado, dosis más bajas del vector extendieron significativamente la supervivencia de ratones nulos para Mecp2 o que presentaban el alelo mutante T158M, aunque no tuvieron impacto en los síntomas neurológicos asociados al RTT. De este modo, se desarrolló un vector AAV9 de segunda generación con un 3'UTR modificado y un promotor extendido que redujo significativamente la toxicidad hepática después de la administración sistémica. Además, la inyección cerebroventricular directa de este vector en ratones neutrales Mecp2 nulos resultó en una alta eficiencia de transducción cerebral, mayor supervivencia y peso corporal, y una mejora de los fenotipos similares a RTT, destacando la importancia de los elementos reguladores endógenos en el casete de expresión génica (Gadalla et al. 2017).

Finalmente, un estudio más reciente diseñó una proteína MeCP2 mínima, que carece de gran parte de ella exceptuando los aminoácidos que codificanpara los dos dominios funcionales de unión a metilo y de interacción NCOR (Tillotson et al. 2017). Cuando ratones neonatales fueron inyectados intracranealmente con un vector scAAV9 que codifica para

esta proteína mínima, mostraron un fenotipo mejorado y supervivencia en ausencia de efectos tóxicos, lo que indica que la función principal de MeCP2 es conectar físicamente el complejo que contiene NCOR al ADN. Este estudio proporciona una nueva vía para realizar estudios de terapia génica, ya que una proteína MeCP2 mínima funcional crea espacio suficiente para secuencias reguladoras adicionales que se pueden empaquetar en la capacidad limitada de los vectores scAAV9. Es importante destacar que esto permitiría un control temporal más preciso de la expresión de Mecp2. Los estudios futuros deberían apuntar a introducir elementos reguladores adicionales en el vector de terapia génica y al mismo tiempo controlar el momento tratamiento para representar mejor la utilidad terapéutica que pueda tener en pacientes humanos.

Un obstáculo adicional del tratamiento de terapia génica para cualquier trastorno implica aumentar la dosis para los humanos. Debido a su pequeño tamaño, los ratones no pueden informar de manera confiable las dosis efectivas para aplicaciones clínicas. Por lo tanto, el tratamiento de terapia génica se está probando en otros modelos animales, incluidos los caninos y los primates no humanos. Los cerebros de los primates no humanos son similares a los humanos con respecto a los circuitos neuronales, la fisiología y las características de comportamiento, lo que los convierte en un modelo ideal para probar la terapia génica para enfermedades neurológicas (Gopinath et al. 2015). Sin embargo, los estudios en modelos más grandes son mucho más caros y requieren períodos experimentales más largos. Es importante destacar que se han generado modelos de primates que carecen de *MECP2* que pueden usarse para acelerar el tratamiento de terapia génica para RTT (Liu et al. 2014; Chen et al. 2017). Recientemente, se ha descubierto que la administración sistémica de dosis altas de AAV9 causa toxicidad hepática y neuronal severa en tres primates no humanos, lo que indica que los esfuerzos en la optimización de la dosis serán imprescindibles (Hinderer et al. 2018).

OBJETIVOS

El objetivo principal de la presente tesis doctoral es la mejora en el diagnóstico genético y la posibilidad de dar un pronóstico sobre desarrollo de la enfermedad en pacientes con RTT/RTT-*like*. Asimismo, profundizar en el estudio de pacientes sin diagnóstico genético para poder determinar la etiología de su clínica.

Los objetivos concretos son:

1. Análisis de variantes en genes relacionados con la clínica RTT y RTT-*like* en una amplia serie de pacientes analizados mediante NGS: valorar las distintas metodologías usadas para el diagnóstico molecular e identificar las causas que puedan explicar el fenotipo de las pacientes tanto en genes conocidos como en genes nuevos.

2. Caracterización clínica y molecular de pacientes con grandes deleciones en *MECP2*: Análisis en profundidad de los mecanismos que conducen a los grandes reordenamientos en *MECP2* e intentar determinar una correlación entre los tamaños de las deleciones y la clínica que desarrollan las pacientes.

3. Análisis de correlaciones genotipo-fenotipo entre RTT y RTT-*like*: estudio de los genes alterados asociados a otras patologías para establecer una posible relación entre pacientes RTT y RTT-*like*, así como estudiar las vías que puedan conectar a los nuevos genes asociados con clínica RTT-*like* a las vías asociadas a *MECP2*.

4. Caracterización funcional de los hallazgos detectados por NGS de mutaciones en genes sin asociación fenotípica que pudieran explicar la clínica de las pacientes.

RESULTADOS

Informe del director de tesis

La memoria de la Tesis Doctoral "Caracterización clínica y molecular del síndrome de Rett: elucidar los casos no resueltos" presentada por Silvia Vidal Falcó, se presenta como un compendio de 6 publicaciones, 5 ya publicadas y una que está en proceso de preparación. La doctoranda Silvia Vidal ha participado de forma activa en la parte experimental, en la interpretación de datos y en la redacción de todos los artículos derivados de esta tesis. Todas las publicaciones que constan aquí forman parte del núcleo de la presente tesis, y no serán utilizadas en otras tesis doctorales, exceptuando el primer artículo que ya ha sido utilizado por Mario Lucariello en su tesis doctoral. Ese primer trabajo fue el artículo seminal de esta tesis, y por eso hemos decidido incluirla en este estudio, ya que fue un trabajo colaborativo entre dos centros de investigación.

PUBLICACIÓN 1. *Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype.* Mario Lucariello, Enrique Vidal, Silvia Vidal, Mauricio Saez, Laura Roa, Dori Huertas, Mercè Pineda, Esther Dalfó, Joaquin Dopazo, Paola Jurado, Judith Armstrong and Manel Esteller. ***Human Genetics*. 2016; 135(12): 1343–1354.** Índice de impacto (2016): 4.637 (Q1)

La doctoranda, Silvia Vidal, ha contribuido en el análisis de datos genómicos, validación por secuenciación Sanger (Figura 1) y la recopilación y análisis de datos clínicos (Tabla 1, 3 y 4). Ha participado en la interpretación de los resultados y en la revisión crítica del manuscrito.

PUBLICACIÓN 2. *The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome.* Silvia Vidal, Núria Brandi, Paola Pacheco, Edgar Gerotina, Laura Blasco, Jean-Rémi Trotta, Sophia Derdak, Maria del Mar O'Callaghan, Àngels Garcia-Cazorla, Mercè Pineda, Judith Armstrong, and *Rett Working Group*. ***Scientific Reports*. 2017 Sep 25;7(1):12288.** Índice de impacto (2017): 4.122 (Q1)

La doctoranda ha intervenido en la realización de este estudio a través de la preparación de librerías del panel *custom* como la de exoma clínico, el estudio de las CNVs a través de los datos genómicos de NGS y la recopilación de datos genéticos y genómicos de pacientes con

RTT estudiados en el Hospital Sant Joan de Déu desde el 1999 hasta el 2016. La doctoranda ha interpretado los datos recopilados y ha realizado estudios estadísticos de los resultados genéticos y comparado las distintas técnicas diagnósticas para este síndrome, redactando el manuscrito e interviniendo en la discusión crítica del mismo.

PUBLICACIÓN 3. *Characterization of large deletions of the MECP2 gene in Rett syndrome patients by gene dosage analysis.* Silvia Vidal, Ainhoa Pascual-Alonso, Marc Rabaza-Gairí, Edgar Gerotina, Núria Brandi, Paola Pacheco, Clara Xiol, Mercè Pineda, *Rett Working Group* and Judith Armstrong. ***Molecular Genetics and Genomic Medicine*. 2019 Aug; 7(8): e793.** Índice de impacto (2018-2019): 2.879 (Q3)

La doctoranda ha sido responsable de gran parte del trabajo experimental, principalmente de los estudios genéticos del gen *MECP2* en DNA genómico (MLPA, Sanger, PCR cuantitativa, PCR largas), que se puede apreciar en la Figura 1 y en la Tabla 1. Ha intervenido en la interpretación de los resultados y en la redacción del manuscrito.

PUBLICACIÓN 4. *The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome.* Silvia Vidal, Núria Brandi, Paola Pacheco, Joan Maynou, Guerau Fernandez, Clara Xiol, Ainhoa Pascual-Alonso, Mercè Pineda, *Rett Working Group* and Judith Armstrong. ***European Journal of Paediatric Neurology*. 2019 Jul;23(4):609-620.** Índice de impacto (2018-2019): 2.496 (Q1)

La doctoranda ha sido la responsable de toda la recopilación de datos clínicos, la selección de pacientes, procedimientos analíticos y del análisis estadístico, y también se ha encargado de la redacción del manuscrito.

PUBLICACIÓN 5. *Genetic Landscape of Rett Syndrome Spectrum: Improvements and Challenges.* Silvia Vidal, Clara Xiol, Ainhoa Pascual-Alonso, del Mar O'Callaghan, Mercè Pineda and Judith Armstrong. ***International Journal of Molecular Sciences*. 2019 Aug 12;20(16).** Índice de impacto (2018-2019): 4.183 (Q1)

La doctoranda ha realizado la recogida de datos clínicos y bibliográficos para la redacción de esta revisión del RTT, que se nos invitó a participar en el 20 aniversario de la descripción de mutaciones en *MECP2* como responsables del RTT. También se ha encargado de la redacción y de la discusión crítica del manuscrito.

PUBLICACIÓN 6. *A de-novo SLC6A1 gene mutation in a patient showing the Rett-like syndrome phenotype.* Silvia Vidal, Mónica Roldan, Mercedes Casado, Janet Hoenicka, Mercè Pineda, Judith Armstrong ***Clinical Genetics (En preparación).***

Este trabajo representa una colaboración con el laboratorio de Genómica Funcional de la Dra. Hoenicka. La doctoranda ha participado en los análisis *in silico* realizados para el estudio de la patogenicidad el cambio estudiado, en el diseño y preparación de los constructos usados para los ensayos *in vitro* de la proteína tanto salvaje como mutada. Asimismo, ha llevado a cabo las transfecciones en los cultivos celulares y los consiguientes ensayos de co-localización de la proteína, su estabilidad con el ensayo de Clorexamida y funcionalidad con el ensayo de captación de GABA. La mutación que se han estudiado proviene del primer trabajo de esta tesis doctoral y se pretendía comprobar la patogenicidad de la variante detectada en el gen *SLC6A1*.

Barcelona, 25 de mayo del 2020

Judith Armstrong Morón

Capítulo 1

Capítulo 1: Estudio molecular del RTT y RTT-*like*. Nuevas aproximaciones metodológicas y aplicaciones diagnósticas.

Hasta el momento hay tres genes descritos y que han sido claramente asociados con el RTT: el gen *MECP2*, descrito como gen principal, y los genes *CDKL5* y *FOXG1*, descritos en un porcentaje menor de casos y asociados a las formas atípicas del RTT. Sin embargo, la etiología de algunos de los pacientes con RTT o con un fenotipo RTT-*like* sigue siendo desconocida. No obstante, con la introducción de las nuevas tecnologías para el diagnóstico genético durante estos últimos años el flujo de trabajo en el laboratorio ha evolucionado de la secuenciación mediante la técnica de Sanger y del análisis de deleciones y duplicaciones mediante MLPA de los tres genes, uno a uno, asociados al RTT al uso de la NGS para el análisis de, a parte de los tres genes, muchos otros que pueden estar implicados en el desarrollo de la enfermedad. Respecto al objetivo 1, se ha analizado una gran cohorte de pacientes RTT haciendo uso de las nuevas tecnologías de secuenciación y se ha procedido a una evaluación de las distintas técnicas, viendo así el rendimiento diagnóstico de estas nuevas metodologías.

Publicación 1

Título: *Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype.*

Autores: Mario Lucariello, Enrique Vidal, Silvia Vidal, Mauricio Saez, Laura Roa, Dori Huertas, Mercè Pineda, Esther Dalfó, Joaquin Dopazo, Paola Jurado, Judith Armstrong and Manel Esteller.

Referencia: *Human Genetics*. 2016; 135(12): 1343–1354.

Resumen:

Basándonos en los pasos previos en la búsqueda de un nuevo gen asociado al RTT que pueda explicar la clínica de las pacientes sin diagnóstico genético y haciendo uso de los avances en el campo de la secuenciación del ADN, se realizó la secuenciación completa del exoma en tríos en una cohorte de 21 pacientes con presentación clínica de RTT siguiendo los criterios establecidos y revisados en 2010 (Neul et al. 2010), pero con diagnóstico genético negativo para los genes *MECP2*, *CDKL5* y *FOXG1*. La hipótesis inicial era la de encontrar un único gen que pudiera explicar la clínica de esta cohorte de pacientes sin diagnóstico genético. Sin embargo, lo que detectamos son mutaciones en múltiples genes candidatos distintos en 14 de los 21 pacientes. Cinco pacientes son portadoras de variantes en genes ya conocidos por estar asociados a otros trastornos sindrómicos del neurodesarrollo. El resto de pacientes con resultado positivo presentaban variantes en genes que en su momento no se habían relacionado todavía con el RTT u otros síndromes del neurodesarrollo. Los genes candidatos fueron validados funcionalmente evaluando la aparición de un fenotipo neurológico en *Caenorhabditis elegans* tras la interrupción del gen ortólogo correspondiente. Estos hallazgos indican que las mutaciones en una gran variedad de genes pueden contribuir al desarrollo de fenotipos similares a RTT, aunque son necesarios más estudios funcionales complementarios que permitan establecer con más fuerza estas predicciones.

Hum Genet (2016) 135:1343–1354
DOI 10.1007/s00439-016-1721-3

ORIGINAL INVESTIGATION

Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype

Mario Lucariello[1] · Enrique Vidal[1] · Silvia Vidal[2] · Mauricio Saez[1] · Laura Roa[1] · Dori Huertas[1] · Mercè Pineda[3] · Esther Dalfó[4] · Joaquin Dopazo[5,6,7] · Paola Jurado[1] · Judith Armstrong[2,8,11] · Manel Esteller[1,9,10]

Received: 15 May 2016 / Accepted: 31 July 2016 / Published online: 19 August 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Classical Rett syndrome (RTT) is a neurodevelopmental disorder where most of cases carry *MECP2* mutations. Atypical RTT variants involve mutations in *CDKL5* and *FOXG1*. However, a subset of RTT patients remains that do not carry any mutation in the described genes. Whole exome sequencing was carried out in a cohort of 21 female probands with clinical features overlapping with those of RTT, but without mutations in the customarily studied genes. Candidates were functionally validated by assessing the appearance of a neurological phenotype in *Caenorhabditis elegans* upon disruption of the corresponding ortholog gene. We detected pathogenic variants that accounted for the RTT-like phenotype in 14 (66.6 %) patients. Five patients were carriers of mutations in genes already known to be associated with other syndromic neurodevelopmental disorders. We determined that the other patients harbored mutations in genes that have not previously been linked to RTT or other neurodevelopmental syndromes, such as the ankyrin repeat containing protein *ANKRD31* or the neuronal acetylcholine receptor subunit alpha-5 (CHRNA5). Furthermore, worm assays demonstrated that mutations in the studied candidate genes caused locomotion defects. Our findings indicate that mutations in a variety of genes contribute to the development of RTT-like phenotypes.

Introduction

Rett syndrome (RTT, MIM 312750) is a postnatal progressive neurodevelopmental disorder (NDD), originally

Electronic supplementary material The online version of this article (doi:10.1007/s00439-016-1721-3) contains supplementary material, which is available to authorized users.

✉ Paola Jurado
pjurado@idibell.cat

✉ Judith Armstrong
jarmstrong@hsjdbcn.org

✉ Manel Esteller
mesteller@idibell.cat

[1] Cancer Epigenetics and Biology Program (PEBC), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet, 08908 Barcelona, Catalonia, Spain

[2] Servei de Medicina Genètica i Molecular, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Esplugues De Llobregat, Catalonia, Spain

[3] Fundació Hospital Sant Joan de Déu (HSJD), Barcelona, Catalonia, Spain

[4] Genetics Department, Bellvitge Biomedical Research Institute (IDIBELL), Barcelona, Catalonia, Spain

[5] Computational Genomics Department, Centro de Investigación Príncipe Felipe (CIPF), 46012 Valencia, Spain

[6] Bioinformatics of Rare Diseases (BIER), CIBER de Enfermedades Raras (CIBERER), Valencia, Spain

[7] Functional Genomics Node (INB) at CIPF, 46012 Valencia, Spain

[8] CIBER Enfermedades Raras, Barcelona, Catalonia, Spain

[9] Department of Physiological Sciences, School of Medicine and Health Sciences, University of Barcelona, Barcelona, Catalonia, Spain

[10] Institucio Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain

[11] Department of Neurology, Hospital Sant Joan de Déu (HSJD), Barcelona, Catalonia, Spain

described in the 1960s by Andreas Rett (Rett 1966), that most frequently manifests itself in girls during early childhood, with an incidence of approximately 1 in 10,000 live births (Chahrour and Zoghbi 2007). RTT patients are asymptomatic during the first 6–18 months of life, but gradually develop severe motor, cognitive, and behavioral abnormalities that persist for life. It is the second most common cause of intellectual disability in females after Down's syndrome (Chahrour and Zoghbi 2007). Around 90 % of the cases are explained by more than 800 reported mutations in the methyl CpG-binding protein 2 gene (*MECP2*) (RettBASE: MECP2 Variation Database) (Christodoulou et al. 2003), which is located in the X chromosome and which causes most of the classical or typical forms of RTT (Chahrour and Zoghbi 2007), and it was originally identified as encoding a protein that binds to methylated DNA (Lewis et al. 1992). Individuals affected by atypical or variant RTT present with many of the clinical features of RTT, but do not necessarily have all of the classic characteristics of the disorder (Neul et al. 2010). Approximately 8 % of classic RTT and 42 % of variant RTT patients are *MECP2* mutation-negative (Monros et al. 2001; Percy 2008). Some of the latter group have mutations in other genes, such as that of the cyclin-dependent kinase-like 5 (*CDKL5*), which is described in individuals with an early seizure onset variant of RTT (Kalscheuer et al. 2003) or the forkhead box G1 (*FOXG1*), which is responsible for the congenital variant of RTT (Ariani et al. 2008). However, there remains a subset of patients with a clinical diagnosis of RTT who are mutation-negative for all the aforementioned genes. Next generation sequencing (NGS) has emerged as a potentially powerful tool for the study of such genetic diseases (Zhu et al. 2015).

Herein, we report the use of a family based exome sequencing approach in a cohort of 20 families with clinical features of RTT, but without mutations in the usually studied genes. We establish the neurological relevance of the newly identified candidate genes by assessing them in *Caenorhabditis elegans* model.

Materials and methods

Patient samples

A cohort of 19 Spanish parent–child trios and one family with two affected daughters who exhibited clinical features associated with RTT were recruited at Sant Joan de Deu Hospital in Barcelona, Catalonia, Spain. These patients had been diagnosed on the basis of the usual clinical parameters (Monros et al. 2001), and according to the recently revised RettSearch International Consortium criteria and nomenclature (Neul et al. 2010), but were found to be mutation-negative for *MECP2*, *CDKL5* and *FOXG1* in the original single-gene screening. The parents were clinically evaluated and it was not observed any evidence of intellectual disability. Genomic DNA from these patients was extracted from peripheral blood leukocytes using standard techniques, and analyzed by exome sequencing at the Cancer Epigenetics and Biology Program (PEBC) in Barcelona, Catalonia, Spain. Ethical approval for the molecular genetic studies was obtained from each institutional review board.

Whole exome sequencing and Sanger validation

Coding regions were captured using the TruSeq DNA Sample Preparation and Exome Enrichment Kit (Illumina, San Diego, California). Paired-end 100 × 2 sequences were sequenced with the Illumina HiScan SQ system at the National Center for Genomic Analysis in Barcelona. We also included the exome sequencing data of an *MECP2*, a *CDKL5* and a *JMJD1C* (Sáez et al. 2016) RTT-associated family for data processing to improve the de novo single nucleotide variant calling. The complete exome sequencing data of all the studied samples are available from the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) with the ID: SRP073424 (private link for the reviewer until publication: http://www.ncbi.nlm.nih.gov/sra/SRP073424). The overall coverage statistics for each individual of the families, considering the regions captures using Exome Enrichment Kit, and number of reads in the position of the variation is shown in Supplementary Table 1. The identified variants were validated by Sanger sequencing using a BigDye® Terminator v3.1 Cycle Sequencing Kit in an Applied Biosystems 3730/DNA Analyzer. The raw data were analyzed with Codon Code Aligner Software. The primers used for Sanger sequencing are shown in Supplementary Table 2.

Caenorhabditis elegans handling

The techniques used for the culture of *Caenorhabditis elegans* were essentially as described (Brenner 1974). The worms were backcrossed at least three times to avoid background mutations. The behavior of three sets of ten animals was independently assessed in locomotion assays without food that were performed at 20 °C, as previously described (Sawin et al. 2000).

Results

Clinical criteria for selecting RTT trios

The 21 patients (derived from the 20 families studied) included in this study fulfilled the recently revised clinical criteria for the diagnosis of RTT following the usual clinical parameters (Monros et al. 2001), and the RettSearch International Consortium criteria and nomenclature (Neul et al. 2010). Specifically, all patients presented stereotypic hand movements, 90.5 % of them (19/21) showed microcephaly and also presented onset of the first signs of the disease before the age of 12 months. 66.7 % of patients (14/21) acquired motor skills, while a further seven (33.3 %), who had a more severe phenotype, never walked. Language skills were progressively lost in 28.6 % of the patients and 71.4 % of them (15/21) never acquired them. Additionally, important episodes of epilepsy were experienced by 81.0 % of the patients (17/21), and 57.1 % of them (12/21) manifested apneas and/or hyperventilation.

Bioinformatic process for filtering and selecting pathogenic variants

Before their inclusion in this study, patients underwent an extensive clinical and genetic work-up to detect genetic alterations in *MECP2*, *CDKL5*, and *FOXG1*. However, no molecular diagnosis could be established. We performed whole exome sequencing (WES) on the 61 individuals (20 pairs of healthy parents and 21 affected daughters) separately by subjecting whole blood derived genomic DNA to exome enrichment and sequencing. We focused our analysis on de novo single nucleotide variants (SNVs) due to their known relevance in autism and mental retardation-related diseases (Vissers et al. 2010). On average, WES gave rise to 419,045 variants, including SNVs and indels, of which 19,951 non-synonymous variants per family (4.7 %) were predicted to have a functional impact on the genomic sequence. To select variants that had not previously been described in the healthy population, we filtered out the variants with an allele frequency of 1 % or higher (the classic definition of a polymorphism) formerly observed in the Single Nucleotide Polymorphism database (dbSNP) and the 1000 Genomes Pilot Project data. Afterwards, to focus on de novo inheritance, patients' variants were filtered first against variants found in their own parents and then against a pool of controls comprising all the healthy parents included in the study. Following this process, we achieved an average of 106 SNVs per family, which corresponded to 81 mutated genes per family. De novo candidate variants were selected on the basis of the quality of the alignments, damage score predictors and the conservation level of each of the genes during evolution. The complete exome sequencing data of all the studied samples are available from the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra).

The global yield of genomic analysis following the bioinformatic process described herein enabled 22 coding de novo mutations to be identified in 66.7 % (14 of 21) of Rett-like patients: 20 SNVs and 2 indels. The identified variants and their de novo status were confirmed by conventional Sanger sequencing. Illustrative samples are shown in Fig. 1. Interestingly, in seven (33.3 %) of the studied RTT probands, exome sequencing did not detect any genetic change relative to their respective parents. The clinical characteristics of these seven patients without obvious pathogenic variants are summarized in Table 1. In one of the families, there were two affected children, and an analysis of potentially relevant recessive variants was performed. For the recessive analysis, and following the same criteria to define a variant as deleterious, we selected the variants with homozygous recessive genotype, and then at the gene level, we also selected the genes presenting more than one heterozygotic variant in the same gene (compound heterozygosity). We did not find any candidate gene consistent with the phenotype of the family with the two affected sisters.

Variants in genes previously associated with neurodevelopmental disorders

Of the 22 identified coding de novo mutations in the assessed RTT-like patients, five (22.7 %) occurred in genes previously associated with neurodevelopmental disorders that presented a clinicopathological phenotype with features coinciding with those of Rett syndrome (Table 2). In particular, we identified four mutations in genes such as *HCN1* (Nava et al. 2014) and *GRIN2B* (Endele et al. 2010; Lemke et al. 2014), which are associated with early infantile epileptic encephalopathy; *SLC6A1*, which is associated with epilepsy and myoclonic-atonic seizures (Carvill et al. 2015); *TCF4*, which is associated with Pitt–Hopkins syndrome (Sweatt 2013); and *SCN1A*, which is associated with Dravet syndrome (Brunklaus and Zuberi 2014) (Table 2). The clinical characteristics of these five patients with variants in genes previously associated with neurodevelopmental phenotypes are summarized in Table 3. A comparison of the clinical features of our RTT-like patients, where we have identified mutations in candidate genes previously associated with other neurodevelopmental disorders, with those observed for these diseases is summarized in Table 4.

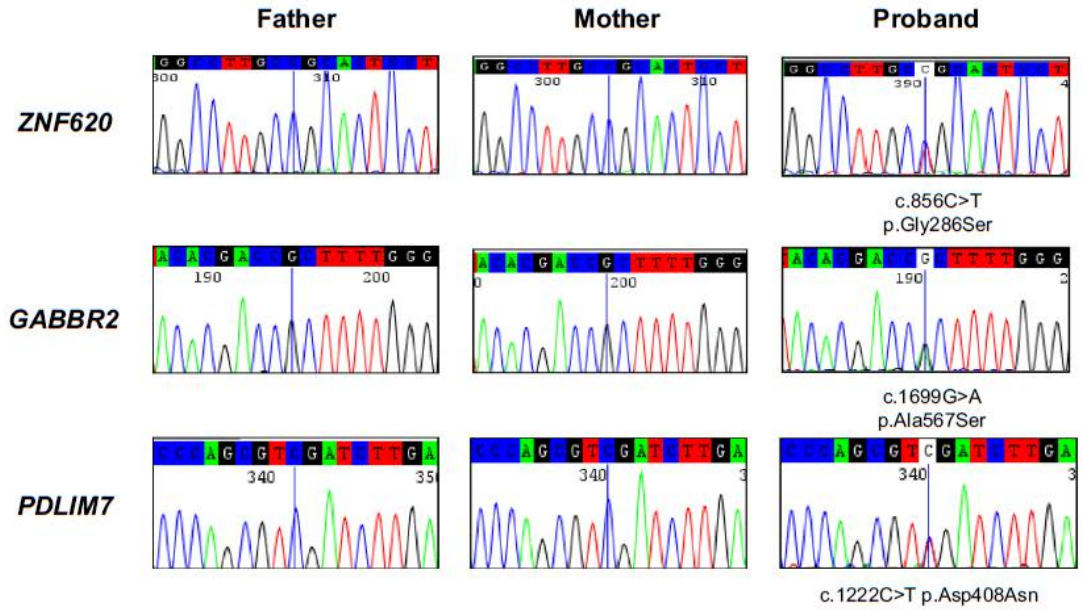


Fig. 1 Sanger sequencing validation of the de novo variants identified by exome sequencing. Illustrative examples for ZNF620 (c.856C > T p.Gly286Ser), GABBR2 (c.1699G > A p.Ala567Ser) and PDLIM7 (c.1222C > T p.Asp408Asn) are shown

Table 1 Clinical summary of patients without exome candidates

Proband	Age (years)	Onset of signs	Micro-cephaly	Sitting alone	Ambula-tion	Respira-tory func-tion	Epilepsy	Hand use	Stereotyp-ies	Language	Total score
1	15	1	1	1	1	0	1	1	1	1	8
2	28	3	1	1	0	1	2	2	3	2	15
5.1	7	3	1	3	4	1	0	3	2	2	19
5.2	5	3	1	1	2	0	0	2	3	2	14
7	16	2	1	0	0	0	2	1	2	1	9
15	8	3	0	1	2	0	2	1	1	1	10
18	3.5	3	1	1	1	0	0	1	3	0	10

Clinical scores of our series of patients according to Pineda scale. Severity classification ranges from 0 to 4 as follows: age of onset of first signs (1: >24 months; 2: 12–24 months; 3: 0–12 months), microcephaly (0: absent; 1: present), sitting alone (1: acquired < 8 months; 2: seat and mantains; 3: seat and lost), ambulation (0: acquired < 18 months, 1: acquired < 30 months; 2: acquired > 30 months; 3: lost acquisition; 4: never acquired), respiratory function (0: no dysfunction; 1: hyperventilation and/or apnea), epilepsy (0: absent; 1: present and controlled; 2: uncontrolled or early epilepsy), hands use (0: acquired and conserved; 1: acquired and partially conserved; 2: acquired and lost; 3: never acquired), onset of stereotypies (1: > 10 years, 2: > 36 months; 3: 18-36 months) and languages (0: preserved and propositive; 1: lost; 2: never acquired). The total score is the sum of the scores of each clinical feature

Variants in genes previously not associated with neurodevelopmental disorders

Of the 22 identified coding de novo variants in the RTT-like patients assessed here, 17 (77.3 %) occurred in genes that had not previously been associated with neurodevelopmental disorders (Table 5). However, two of these variants were associated with non-neurodevelopmental disorders: a *BTBD9* variant linked to restless leg syndrome (Kemlink et al. 2009), and an *ATP8B1* SNV associated with familial cholestasis (Klomp et al. 2004), respectively. Interestingly, the *BTBD9* variant was detected in the same patient that carried the *SCN1A* variant associated with Dravet syndrome (Table 2). The other 15 potentially pathogenic variants identified occurred in genes that had not been linked to any genetic disorder of any type. However, there was an enrichment of genes with a potential role in neuronal biology and functionality, such as the gamma-aminobutyric type

Table 2 List of patients with variants found in genes previously associated with neurodevelopmental phenotypes

Proband	Gene	Protein	NM number	Variant: genomic coordinates	cDNA change	Protein change	Gene-disease association
4	*HCN1*	Hyperpolarization Activated Cyclic Nucleotide Gated Potassium Channel 1	NM_021072.3	5:45396665	c.1159G > T	p.Ala387Ser	Early infantile epileptic encephalopathy 24
8	*SCN1A*	Sodium Channel Protein Type I Subunit Alpha	NM_001165963.1	2:166866266	c.3965C > G	p.Arg1322Thr	Dravet syndrome
10	*TCF4*	Transcription Factor 4	NM_001243236.1	18:52901827	c.958delC	p.Gln320Ser_fs8X	Pitt–Hopkins syndrome
11	*GRIN2B*	Glutamate receptor ionotropic, NMDA 2B	NM_000834.3	12:13764782	c.1657C > A	p.Pro553Thr	Autosomal Dominant Mental Retardation 6; Early infantile epileptic encephalopathy 27
17	*SLC6A1*	Solute Carrier Family 6 Member 1	NM_003042.3	3:11067528	c.919G > A	p.Gly307Arg	Myoclonic-atonic epilepsy and schizophrenia

B receptor subunit 2 (*GABBR2*), the neuronal acetylcholine receptor subunit alpha-5 (*CHRNA5*), the Huntington-associated protein 1 (*HAP1*), the axon guider semaphorin 6B, the ankyrin repeat containing proteins *ANKRD31* and *AGAP6*, and the neuronal voltage-gated calcium channel *CACNA1* (Table 5). Proband 14 was a particularly interesting case in which four potential pathogenic variants were present, affecting zinc finger (*ZNF620*), a nucleolar complex (*NOC3L*), G patch domain (*GPATCH2*) and GRAM domain (*GRAMD1A*)-related proteins (Table 5). The clinical characteristics of these patients with variants in genes previously not associated with neurodevelopmental disorders are summarized in Table 6.

Neurological phenotype of candidate genes in *C. elegans*

To demonstrate a neurological effect for a loss of function of the detected genes that had not previously been associated with neurodevelopmental disorders (Table 5), we used the model organism *C. elegans* to confirm the genotype-phenotype correlation. We obtained all the available *C. elegans* mutants that carry deleterious mutations in the orthologous genes to those human genes with potentially pathogenic mutations in the patients. In this model, backcrossing is a commonly used procedure to obtain a specific mutant strain without any secondary mutations from its genetic composition. Under these conditions, we were able to test six available mutant strains that were backcrossed at least three times to prove that any observed phenotype was really associated to specific mutations in the orthologous genes. To this end, we studied the *C. elegans* mutants carrying deleterious mutations in the gene orthologs of the human genes *PDLIM7*, *ANKRD31*, *ZNF620*, *CHRNA5*, *MGRN1* and *GABBR2* described in Table 7. Considering that the loss of normal movement and coordination is one of the clearest signs shown by Rett patients, we performed a locomotion assay of the nematodes as previously described (Sawin et al. 2000), using the wild-type N2 strain as a control (Supplementary Video 1). We observed that in 83.3 % (5 of 6) of the cases the mutation of the ortholog of the human exome sequencing identified genes in *C. elegans* exhibited a locomotion defective phenotype (Fig. 2). The most severe phenotypes were represented by *alp-1*, *unc-44* and *pag-3*, with mutations in the orthologs of PDZ and LIM domain protein 7 (*PDLIM7*), ankyrin repeat containing protein *ANKRD31* and the zinc protein *ZNF620*, respectively (Fig. 2 and Supplementary Videos 2, 3 and 4). The case of *alp-1* was particularly interesting, because mutant worms were not only thinner than usual and completely locomotion defective, but they exhibited transitory spasms. Significant defects, such as slower locomotion and uncoordinated movement, were also observed in the mutants of *unc-63* and *C11H1.3*, the *C. elegans* orthologs

Table 3 Clinical summary of the patients with variants in genes previously associated with neurodevelopmental phenotypes

Proband	Gene variant	Age (years)	Onset of the signs	Microcephaly	Sitting alone	Ambulation	Respiratory function	Epilepsy	Hand use	Stereotypies	Language	Total score
4	*HCN1*	24	3	1	1	3	1	2	3	1	2	16
8	*SCN1A, MGRN1, BTBD9*	7	3	1	1	4	1	2	1	3	2	18
10	*TCF4*	16	3	1	1	2	1	2	1	1	2	14
11	*GRIN2B, SEMA6B*	3	2	0	1	3	0	1	1	1	2	10
17	*SLC6A1*	36	3	0	1	1	0	1	1	1	0	7

Clinical scores of our series of patients according to Pineda scale. Severity classification ranges from 0 to 4 as follows: age of onset of first signs (1: > 24 months; 2: 12–24 months; 3: 0–12 months), microcephaly (0: absent; 1: present), sitting alone (1: acquired < 8 months; 2: seat and mantains; 3: seat and lost), ambulation (0: acquired < 18 months, 1: acquired < 30 months; 2: acquired > 30 months; 3: lost acquisition; 4: never acquired), respiratory function (0: no dysfunction; 1: hyperventilation and/or apnea), epilepsy (0: absent; 1: present and controlled; 2: uncontrolled or early epilepsy), hands use (0: acquired and conserved; 1: acquired and partially conserved; 2: acquired and lost; 3: never acquired), onset of stereotypies (1: > 10 years, 2: > 36 months; 3: 18-36 months) and languages (0: preserved and propositive; 1: lost; 2: never acquired). The total score is the sum of the scores of each clinical feature

of the genes coding for the neuronal acetylcholine receptor subunit alpha-5 (*CHRNA5*) and mahogunin RING finger protein 1 (*MGRN1*), respectively. Although we did not find a clear locomotion defect in the *gbb-2* mutant (the ortholog of *GABBR2*) (Fig. 2), it occurs in the *gbb-1*;*gbb-2* double mutant (Dittman and Kaplan 2008), *gbb-1* being the *C. elegans* ortholog of *GABBR1* (gamma-aminobutyric acid type B receptor subunit 1). The clinical picture of the particular RTT cases with mutations in the genes studied in *C. elegans* is shown in Table 6.

Discussion

Our results indicate that the existence of de novo variants in genes with potential neurological functionalities, such as neuronal receptors (*GABBR2* and *CHRNA5*), axon guiders (*SEMA6B*), synaptic ionic channels (*CACNA1I*) and others, contribute to the development of RTT-like clinical phenotypes in the context of wild-type sequences for standard Rett genes such as *MECP2* and *FOXG1*. These patients share most of the clinicopathological features of classic RTT syndrome, such as stereotypic hand movements, relative microcephaly, and onset of the disease after the age of 12 months. Thus, exome sequencing is a powerful tool for genetically characterizing these enigmatic cases. In this regard, once a new candidate gene has been identified, it is now possible to design specific sequencing strategies for the molecular screening of this particular target in larger populations of patients with intellectual disability. The strategy based on exome sequencing patients who have RTT features, but no known mutations in the usual genes, has recently been used in other smaller series of patients (Grillo et al. 2013; Okamoto et al. 2015; Hara et al. 2015; Olson et al. 2015; Lopes et al. 2016). Most importantly, our study and the aforementioned previous reports strengthen the concept that a mutational heterogeneous profile hitting shared neurological signaling pathways contributes to RTT-like syndromes. Examples of confluence in the same molecular crossroads include the gamma-aminobutyric type B receptor subunit 2 (*GABBR2*) de novo variant, described here, and the formerly identified variant in the gamma-aminobutyric acid receptor delta gene (*GABRD*) (Hara et al. 2015). Interestingly, a second RTT-like patient has been identified as being a carrier of a de novo *GABBR2* variant (Lopes et al. 2016), highlighting the likelihood that this gene and pathway contribute to the clinical entity. Another example of similarly targeted genes in RTT-like patients is that of the proteins containing ankyrin-repeats that are involved in postsynaptic density (Durand et al. 2007). This study has revealed de novo variants in the ankyrin repeat containing proteins *AGAP6* and *ANKRD31* in RTT-like patients, and the presence of de novo variant of the SH3 and multiple ankyrin repeat domain3 protein (*SHANK3*) (Hara et al. 2015) and ankyrin-3 (*ANK3*) (Grillo et al. 2013) has been reported in two RTT-like patients. A final example of the convergence of cellular pathways to provide a common RTT-like phenotype is represented by the disruption of the ionic channels. We found the existence of a voltage-gated calcium channel subunit alpha 11 (*CANA1I*) de novo variant in an RTT-like patient. Additionally, the presence of de novo variants in the calcium release channel *RYR1* (Grillo et al. 2013) and the sodium voltage-gated channel alpha subunit 2 (*SCN2A*) (Baasch et al. 2014) in two other RTT-like

Table 4 Comparison of the clinical features of RTT-like patients from who we have identified mutations in candidate genes previously associated with other neurodevelopmental disorders with those observed for these diseases

Disease	Rett		Atypical Rett		Pitt–Hopkins		Dravet		EEIE27		MAE		EEIE24	
GENE	*MECP2*		*CDKL5*		*TCF4*		*SCN1A*		*GRIN2B*		*SLC6A1*		*HCN1*	
OMIM/Patient	312,750	6	308,350	16	602,272	10	182,390	8	616,139	11	616,421	17	615,871	4
Onset age	6–18 m	12–24 m	1–3 m	0–12 m	6–12 m	0–12 m	2 days–7 m	0–12 m	0–24 m	12–24 m	0–24 m	0–12 m	0–24 m	0–12 m
Microcephaly	Yes	Yes	Yes	Yes	Yes	Yes	±	Yes	±	No	±	No	NA	Yes
Hypotonia	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	±	Yes	±	Yes	NA	Yes
Epilepsy	80 %	Yes	Yes	Yes	Yes	Yes	Yes	Yes	±	Yes	Yes	Yes	Yes	Yes
Respiratory dysfunction	80 %	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	NA	Yes
Expressive language dysfunction	Yes	Yes	Yes	Yes	Yes	Yes	±	Yes	Yes	Yes	No	No	NA	Yes
Preserved use of hands	No	No	±	No	No	No	Yes	No	±	No	±	No	NA	No
Stereotypies	Yes	Yes	Yes	Yes	Yes	Yes	±	Yes	±	Yes	±	Yes	NA	Yes
Inheritance	XL		XL		AD		AD		AD		AD		AD	

EEIE epileptic encephalopathy, early infantile, *MAE*, myoclonic-atonic epilepsy, *m* months, *XL* X linkage, *AD* autosomal dominant, *NA* unavailable

probands have been reported. It is also intriguing that in our study a variant in HAP was found, whereas in similar series heterozygous variants in huntingtin (HTT) have been described (Lopes et al. 2016; Rodan et al. 2016), further reinforcing the links between Huntington's disease and Rett syndrome (Roux et al. 2012). Another interesting case is provided by TCF4, which is associated with Pitt–Hopkins syndrome (Sweatt 2013), where in addition to our study, others have found mutations in RTT-like patients (Lopes et al. 2016). This observation could be of interest for clinicians due to phenotypic similitudes such as intellectual disability, stereotypic movement, apneas and seizures (Marangi et al. 2012).

Our findings also suggest that a substantial degree of clinical overlap can exist between the features associated with RTT and those of other neurodevelopmental disorders. Our exome sequencing effort indicated that probands originally diagnosed as RTT-like patients were, in fact, carriers of well-known pathogenic de novo mutations linked to Dravet Syndrome (*SCN1A*), myoclonic-atonic epilepsy (*SCLC6A1*), or early infantile epileptic encephalopathies 24 (*HCN1*) and 27 (*GRIN2B*). The purely clinical classification of these patients, without a thorough genetic study, can be difficult because some of these patients are composites that carry at least two pathogenic variants. For example, in our cases, the Dravet syndrome patient also had a de novo variant in *BTBD9* associated with the development of restless leg syndrome. In addition, among the newly identified candidate genes associated with RTT-like features, a few of these patients simultaneously carried two de novo variants (e.g., probands 8, 19 and 21), further complicating the tasks of correctly diagnosing and managing these individuals.

Finally, the studies performed in *C. elegans* validate the functional relevance for nervous system function of the newly proposed candidate genes. Future studies would be necessary to assess the role of the specific variants identified, such as rescuing the defects with the expression of normal cDNAs versus cDNAs containing the mutation, ideally using cDNAs of human origin to prove similar function of the gene in the two species. It is also relevant to mention that for some of the newly reported mutated genes in our RTT-like patients, there are mice models targeting the described loci that show neurological phenotypes such as BTBD9 (motor restlessness and sleep disturbances) (DeAndrade et al. 2012), MGRN1 (spongiform neurodegeneration) (He et al. 2003), SEMA6B (aberrant mossy fibers) (Tawarayama et al. 2010), CHRNA5 (alterations in the habenulo-interpeduncular pathway) (Fowler et al. 2011), GABBR2 (anxiety and depression-related behavior) (Mombereau et al. 2005) and HAP1 (depressive-like behavior and reduced hippocampal neurogenesis) (Chan et al. 2002; Xiang et al. 2015).

publi1.oxps - Visor de XPS

Table 5 List of patients with variants in new candidate disease genes

Proband	Gene	Protein	Function	NM number	Variant: genomic coordinates	cDNA change	Protein change	ExAC	SIFT	Polyphen2	PROVEAN	Mutation Taster2	Conservation
3	*AGAP6*	ArfGAP with GTPase domain, ankyrin repeat and PH domain 6	Putative GTPase-activating protein	NM_001077665.2	10:51748528	c.53insC	p.Asp18Ala_fs10X	Not present	NA	NA	B	P	405
8	*MGRN1*	Mahogunin RING Finger Protein 1	E3 ubiquitin-protein ligase	NM_001142290.2	16:4723583	c.880C > T	p.Arg294Cys	0.000077	P	P	P	P	573
8	*BTBD9*	BTB (POZ) Domain-Containing 9	Putative protein–protein interactor	NM_001099272.1	6:38256093	c.1409C > T	p.Ala470Val	Not present	B	P	P	P	512
11	*SEMA6B*	Semaphorin-6B	Role in axon guidance	NM_032108.3	19:4555540	c.508G > A	p.Gly170Ser	Not present	P	P	P	P	510
12	*VASH2*	Vasohibin 2	Angiogenesis inhibitor	NM_001301056.1	1:213161902	c.1044A > C	p.Glu348Asp	Not present	B	B	B	B	473
13	*CHRNA5*	Neuronal acetylcholine receptor subunit alpha-5	Excitator of neuronal activity	NM_000745.3	15:78882481	c.748C > A	p.Pro250Thr	Not present	B	P	P	P	519
14	*ZNF620*	Zinc Finger Protein 620	Transcriptional regulator	NM_175888.3	3:40557941	c.856G > A	p.Gly286Ser	Not present	P	P	P	P	317
14	*GRAMD1A*	GRAM Domain-Containing 1A	Not described	NM_020895.3	19:35506764	c.1106G > A	p.Arg369His	Not present	P	P	P	P	358
14	*NOC3L*	Nucleolar complex protein 3 homolog	Regulator of adipogenesis	NM_022451.10	10:96097586	c.2137G > A	p.Ala713Thr	Not present	B	B	B	B	0
14	*GPATCH2*	G patch domain-containing protein 2	Regulator of cell proliferation	NM_018040.3	1:217784371	c.878G > A	p.Gly293Asp	Not present	B	P	P	P	304

Table 5 continued

Proband	Gene	Protein	Function	NM number	Variant: genomic coordinates	cDNA change	Protein change	ExAC	SIFT	Polyphen2	PROVEAN	Mutation Taster2	Conservation
19	*GABBR2*	Gamma-aminobutyric acid type B receptor subunit 2	Inhibitor of neuronal activity	NM_005458.7	9:101133817	c.1699G > A	p.Ala567Thr	Not present	P	P	P	P	412
19	*ATP8B1*	Phospholipid-transporting ATPase IC	Aminophospholipid translocator	NM_005603.4	18:55328507	c.2606C > T	p.Thr869Ile	Not present	P	P	P	P	361
20	*HAP1*	Huntingtin-Associated Protein 1	Vesicular transporter	NM_177977.2	17:39890655	c.232G > A	p.Ala78Thre	Not present	P	B	B	B	0
21	*PDLIM7*	PDZ and LIM domain protein 7	Scaffold protein	NM_005451.4	5:176910933	c.1222G > A	p.Asp408Asn	Not present	P	P	B	P	515
21	*SRRM3*	Serine/Arginine Repetitive Matrix 3	Splicing activator	NM_001291831.1	7:75890878	c.655C > G	p.Ser218Cys	Not present	P	P	P	P	491
22	*ANKRD31*	Ankyrin Repeat Domain 31	Not described	NM_001164443.1	5:74518166	c.196A > T	p.Ile66Phe	Not present	P	P	B	B	401
23	*CACNA1I*	Voltage-Gated Calcium Channel Subunit Alpha 1I	Calcium signaling in neurons	NM_021096.3	22:40066855	c.4435C > T	p.Leu1479Phe	Not present	B	P	B	P	695

ExAC, frequency of the identified variants in the exome aggregation consortium. Four in silico prediction tools of functional mutation impact were used: 'Sorting Tolerant From Intolerant' (*SIFT*), 'Polymorphism Phenotyping v2' (*Polyphen2*); 'Protein Variation Effect Analyzer' (*PROVEAN*) and Mutation Taster2. The output results were classified as: likely pathogenic (*P*), likely benign (*B*) and not available (*NA*). Conservation scores refer to the conservation level of the nucleotide at the position of the identified variant between 46 species of vertebrates based on PhastCons. It ranges from 0 to 1000: the highest, the more conserved during evolution

Table 6 Clinical summary of patients with variants in new candidate disease genes

Proband	Gene variant	Age (years)	Onset of the signs	Micro-cephaly	Sitting alone	Ambu-lation	Res-piratory function	Epilepsy	Hand use	Stereo-typies	Lan-guage	Total score
3	*AGAP6*	14	3	1	2	4	1	1	3	2	2	19
8	*SCN1A, MGRN1, BTBD9*	7	3	1	1	4	1	2	1	3	2	18
11	GRIN2B, SEMA6B	3	2	0	1	3	0	1	1	1	2	10
12	*VASH2*	11	3	1	1	4	1	0	2	2	1	15
13	*CHRNA5*	10	3	1	2	4	1	1	2	3	2	19
14	ZNF620, GRAMD1A, NOC3L, GPATCH2	2	3	1	1	2	0	1	3	2	2	15
19	GABBR2, ATP8B1	2	3	1	1	4	0	0	3	2	2	16
20	*HAP1*	24	3	1	1	1	1	0	2	2	1	12
21	*PDLIM7, SRRM3*	5	3	1	3	4	0	1	2	1	2	17
22	*ANKRD31*	17	3	0	1	2	0	1	2	3	2	14
23	*CACNA1I*	1/8	3	1	1	2	1	0	3	3	2	16

Clinical scores of our series of patients according to Pineda scale. Severity classification ranges from 0 to 4 as follows: age of onset of first signs (1: >24 months; 2: 12–24 months; 3: 0–12 months), microcephaly (0: absent; 1: present), sitting alone (1: acquired < 8 months; 2: seat and mantains; 3: seat and lost), ambulation (0: acquired < 18 months, 1: acquired < 30 months; 2: acquired > 30 months; 3: lost acquisition; 4: never acquired), respiratory function (0: no dysfunction; 1: hyperventilation and/or apnea), epilepsy (0: absent; 1: present and controlled; 2: uncontrolled or early epilepsy), hands use (0: acquired and conserved; 1: acquired and partially conserved; 2: acquired and lost; 3: never acquired), onset of stereotypies (1: > 10 years, 2: > 36 months; 3: 18–36 months) and languages (0: preserved and propositive; 1: lost; 2: never acquired). The total score is the sum of the scores of each clinical feature

Table 7 Phenotype in *Caenorhabditis elegans*

Human gene	Ortholog in C.elegans	Similarity (%)	Identity (%)	Mutation in C.elegans	Locomotion phenotype	Neurological phenotypes	Other phenotypes
GABBR2	*gbb-2*	53	34	Deletion	normal	Hypersensitivity to aldicarb	–
MGRN1	*C11H1.3*	58	41	Deletion	locomotion defective	–	–
CHRNA5	*unc-63*	58	40	Deletion	locomotion defective	Uncoordinated locomotion with strong levamisole resistance	–
ZNF620	*pag-3*	65	47	Deletion	locomotion defective	Altered neurosecretion and up-regulation of DCV (Dense Core Vesicles) components	–
ANKRD31	*unc-44*	59	39	Deletion	locomotion defective	Asymmetric dynamics of axonal and dendritic microtubules defects	–
PDLIM7	*alp-1*	65	47	Deletion	locomotion defective	–	Defects in actin filament organization in muscle cells

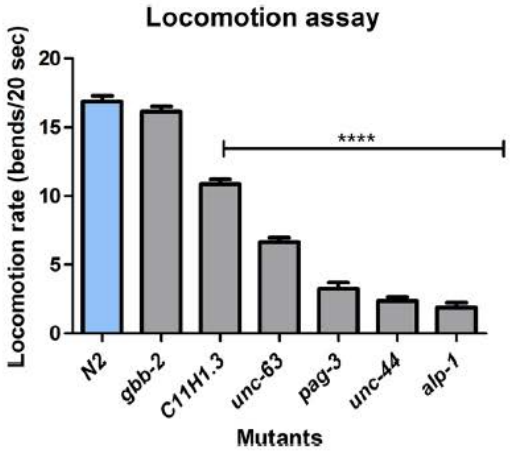


Fig. 2 Locomotion assay in *Caenorhabditis elegans*. Functional validation of mutations was performed by measuring the locomotion rate, expressed in average of measuring, *in C. elegans*. Each mutant strain was compared to a wild-type N2 control strain by measuring worm body bends during 20 s in three independent sets of experiments. Locomotion rates of mutants, represented by *C11H1.3* (*MGRN1*), *unc-63* (*CHRNA5*), *pag-3* (*ZNF620*), *unc-44* (*ANKRD31*) and *alp-1* (*PDLIM7*) are significantly lower compared to that of the N2 control strain ($p < 0.0001$), on the contrary *gbb-2* (*GABBR2*) mutant move similarly. Standard error of the mean (SEM) values is shown. *p* values obtained according to Student's *t* test. ***$p < 0.0001$

Conclusions

Overall, this study demonstrates the genetic mutational diversity that underlies the clinical diagnosis of patients with clinical features that resemble RTT cases. Once the recognized *MECP2*, *CDKL5* and *FOXG1* mutations have been discarded, exome sequencing emerges as a very useful strategy for the more accurate classification of these patients. The de novo variants identified by this approach can modify the first diagnostic orientation towards another neurodevelopmental disorder, or pinpoint new genes involved in the onset of RTT-like features. Interestingly, most of these new targets are involved in the same functional networks associated with correct neuronal functionality. Further research is required to understand the role of these proteins in the occurrence of neurodevelopmental diseases. Additional functional experiments, such as the *C. elegans* assays used in this study, would be extremely helpful for this purpose.

Acknowledgments The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2012 under REA Grant agreement PITN-GA-2012-316758 of the EPITRAIN project and PITN-GA-2009–238242 of DISCHROM; the E-RARE EuroRETT network (Carlos III Health Institute project PI071327); the Foundation Lejeune (France); the Cellex Foundation; the Botín Foundation; the Finestrelles Foundation; the Catalan Association for Rett Syndrome; and the Health and Science Departments of the Catalan government (Generalitat de Catalunya). M.E. is an ICREA Research Professor.

Compliance with ethical statement

Conflict of interest The authors declare that there is no conflict of interest associated with this manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Ariani F et al (2008) FOXG1 is responsible for the congenital variant of Rett syndrome. Am J Hum Genet 83:89–93

Baasch AL et al (2014) Exome sequencing identifies a de novo SCN2A mutation in a patient with intractable seizures, severe intellectual disability, optic atrophy, muscular hypotonia, and brain abnormalities. Epilepsia 55:e25–e29

Brenner S (1974) The genetics of Caenorhabditis elegans. Genetics 77:71–94

Brunklaus A, Zuberi SM (2014) Dravet syndrome–from epileptic encephalopathy to channelopathy. Epilepsia 55:979–984

Carvill GL et al (2015) Mutations in the GABA transporter SLC6A1 cause epilepsy with myoclonic-atonic seizures. Am J Hum Genet 96:808–815

Chahrour M, Zoghbi HY (2007) The story of Rett syndrome: from clinic to neurobiology. Neuron 56:422–437

Chan EY et al (2002) Targeted disruption of Huntingtin-associated protein-1 (Hap1) results in postnatal death due to depressed feeding behavior. Hum Mol Genet 11:945–959

Christodoulou J, Grimm A, Maher T, Bennetts B (2003) RettBASE: the IRSA MECP2 Variation Database—a new mutation database in evolution. Hum Mutat 21:466–472

DeAndrade MP et al (2012) Motor restlessness, sleep disturbances, thermal sensory alterations and elevated serum iron levels in Btbd9 mutant mice. Hum Mol Genet 21:3984–3992

Dittman JS, Kaplan JM (2008) Behavioral impact of neurotransmitter-activated G-protein-coupled receptors: muscarinic and GABAB receptors regulate Caenorhabditis elegans locomotion. J Neurosci 28:7104–7112

Durand CM et al (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat Genet 39:25–27

Endele S et al (2010) Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. Nat Genet 42:1021–1026

Fowler CD, Lu Q, Johnson PM, Marks MJ, Kenny PJ (2011) Habenular α5 nicotinic receptor subunit signalling controls nicotine intake. Nature 471:597–601

Grillo E et al (2013) Revealing the complexity of a monogenic disease: Rett syndrome exome sequencing. PLoS One 8:e56599

Hara M, Ohba C, Yamashita Y, Saitsu H, Matsumoto N, Matsuishi T (2015) De novo SHANK3 mutation causes Rett syndrome-like phenotype in a female patient. Am J Med Genet A 167:1593–1596

He L et al (2003) Spongiform degeneration in mahoganoid mutant mice. Science 299:710–712
Kalscheuer VM et al (2003) Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. Am J Hum Genet 72:1401–1411
Kemlink D et al (2009) Replication of restless legs syndrome loci in three European populations. J Med Genet 46:315–318
Klomp LW et al (2004) Characterization of mutations in ATP8B1 associated with hereditary cholestasis. Hepatology 40:27–38
Lemke JR et al (2014) GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. Ann Neurol 75:147–154
Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, Bird A (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 69:905–914
Lopes F et al (2016) Identification of novel genetic causes of Rett syndrome-like phenotypes. J Med Genet 53:190–199
Marangi G et al (2012) Proposal of a clinical score for the molecular test for Pitt-Hopkins syndrome. Am J Med Genet A 158A:1604–1611
Mombereau C, Kaupmann K, Gassmann M, Bettler B, van der Putten H, Cryan JF (2005) Altered anxiety and depression-related behaviour in mice lacking GABAB(2) receptor subunits. NeuroReport 16:307–310
Monrós E, Armstrong J, Aibar E, Poo P, Canós I, Pineda M (2001) Rett syndrome in Spain: mutation analysis and clinical correlations. Brain Dev 23:S251–S253
Nava C et al (2014) De novo mutations in HCN1 cause early infantile epileptic encephalopathy. Nat Genet 46:640–645
Neul JL et al (2010) Rett syndrome: revised diagnostic criteria and nomenclature. Ann Neurol 68:944–950
Okamoto N et al (2015) Targeted next-generation sequencing in the diagnosis of neurodevelopmental disorders. Clin Genet 88:288–292
Olson HE et al (2015) Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. Am J Med Genet A 167A:2017–2025
Percy AK (2008) Rett syndrome: recent research progress. J Child Neurol 23:543–549
Rett A (1966) On an unusual brain atrophy syndrome in hyperammonemia in childhood. Wien Med Wochenschr 116:723–772
Rodan LH et al (2016) A novel neurodevelopmental disorder associated with compound heterozygous variants in the huntingtin gene. Eur J Hum Genet. doi:10.1038/ejhg.2016.74
Roux JC, Zala D, Panayotis N, Borges-Correia A, Saudou F, Villard L (2012) Modification of Mecp2 dosage alters axonal transport through the Huntingtin/Hap1 pathway. Neurobiol Dis 45:786–795
Sáez MA et al (2016) Mutations in JMJD1C are involved in Rett syndrome and intellectual disability. Genet Med 18:378–385
Sawin ER, Ranganathan R, Horvitz HR (2000) C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway. Neuron 26:619–631
Sweatt JD (2013) Pitt-Hopkins syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. Exp Mol Med 45:e21
Tawarayama H, Yoshida Y, Suto F, Mitchell KJ, Fujisawa H (2010) Roles of semaphorin-6B and plexin-A2 in lamina-restricted projection of hippocampal mossy fibers. J Neurosci 30:7049–7060
Vissers LE et al (2010) A de novo paradigm for mental retardation. Nat Genet 42:1109–1112
Xiang J, Yan S, Li SH, Li XJ (2015) Postnatal loss of hap1 reduces hippocampal neurogenesis and causes adult depressive-like behavior in mice. PLoS Genet 11:e1005175
Zhu X et al (2015) Whole-exome sequencing in undiagnosed genetic disorders: interpreting 119 trios. Genet Med 17:774–781

Publicación 2

Título: *The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome.*

Autores: Silvia Vidal, Núria Brandi, Paola Pacheco, Edgar Gerotina, Laura Blasco, Jean-Rémi Trotta, Sophia Derdak, Maria del Mar O'Callaghan, Àngels Garcia-Cazorla, Mercè Pineda, Judith Armstrong, and *Rett Working Group*.

Referencia: *Scientific Reports*. 2017 Sep 25;7(1):12288

Resumen:

Recientemente, la NGS ha impulsado una mejora en el rendimiento del diagnóstico genético debido a la rapidez y la asequibilidad del método. Para evaluar la utilidad de estos métodos en el diagnóstico genético, presentamos este estudio retrospectivo donde comparamos los resultados genéticos obtenidos en pacientes RTT/RTT-*like* en el que se han utilizado diferentes técnicas basadas en estas nuevas tecnologías y en la metodología previa, la secuenciación Sanger. En total, hemos analizado nuestra cohorte completa de pacientes con diagnósticos clínicos de RTT y RTT-*like* (1577) y reanalizado aquellos casos que tenían resultados negativos para la secuenciación Sanger de *MECP2*, *CDKL5* y *FOXG1*. Las metodologías basadas en la NGS empleadas han sido: panel prediseñado de 17 genes asociados a clínica RTT-*like*, exoma clínico donde se incluyen los genes que han sido asociados a clínica y WES donde se analizan todas las regiones codificantes del genoma humano. De este modo, con este estudio hemos podido concluir que la mejor opción para empezar el diagnóstico genético de nuevos casos en nuestro laboratorio es el exoma clínico, ya que nos da opción a agrupar en un mismo ensayos muestras de pacientes de diversas enfermedades y nos proporciona los parámetros de calidad correctos, marcados por los estándares de calidad. Por el contrario, aquellos casos con diagnóstico genético previo negativo la mejor opción es realizar un ensayo WES junto con las muestras de los progenitores. Finalmente, podemos concluir que la NGS ha mejorado nuestra capacidad de dar diagnóstico genético a las pacientes.

www.nature.com/scientificreports

OPEN

The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome

Received: 23 June 2017
Accepted: 29 August 2017
Published online: 25 September 2017

Silvia Vidal[1], Núria Brandi[2], Paola Pacheco[1], Edgar Gerotina[1], Laura Blasco[1], Jean-Rémi Trotta[3], Sophia Derdak[3], Maria del Mar O'Callaghan[4,5,6], Àngels Garcia-Cazorla[4,5,6], Mercè Pineda[5], Judith Armstrong[1,5,6] & Rett Working Group*

Rett syndrome (RTT) is an early-onset neurodevelopmental disorder that almost exclusively affects girls and is totally disabling. Three genes have been identified that cause RTT: *MECP2*, *CDKL5* and *FOXG1*. However, the etiology of some of RTT patients still remains unknown. Recently, next generation sequencing (NGS) has promoted genetic diagnoses because of the quickness and affordability of the method. To evaluate the usefulness of NGS in genetic diagnosis, we present the genetic study of RTT-like patients using different techniques based on this technology. We studied 1577 patients with RTT-like clinical diagnoses and reviewed patients who were previously studied and thought to have RTT genes by Sanger sequencing. Genetically, 477 of 1577 patients with a RTT-like suspicion have been diagnosed. Positive results were found in 30% by Sanger sequencing, 23% with a custom panel, 24% with a commercial panel and 32% with whole exome sequencing. A genetic study using NGS allows the study of a larger number of genes associated with RTT-like symptoms simultaneously, providing genetic study of a wider group of patients as well as significantly reducing the response time and cost of the study.

Rett syndrome (RTT; MIM# 312750) is a neurodevelopmental disorder of early onset that affects girls almost exclusively. RTT was originally described in the 1960s by Andreas Rett[1]. This syndrome is first recognized in infancy with a period of apparently normal development (up to the age of 6–18 months), followed by a regression characterized by loss of speech and purposeful hand use and motor apraxia that may be associated with epilepsy and dysautonomic features, including disturbed breathing, sleep and gastrointestinal motility[1,2]. RTT has an incidence of 1:10,000 live female births and is the second cause of intellectual disability after Down's syndrome in females[3]. RTT was clinically cataloged into classic and atypical forms of the disease. However, these criteria have undergone several updates over the past three decades. Consensus criteria have been established that distinguish RTT patients into the individual classifications of classic or typical RTT and the atypical or variant forms of RTT[4–6].

A large number of reports support the evidence that mutations in the Methyl CpG binding protein 2 gene (*MECP2*; MIM *300005) are the major causes of classical RTT[7,8]. Over 95% of the cases are explained by more than 800 reported mutations in the methyl CpG-binding protein 2 gene (*MECP2*) (RettBASE: MECP2 Variation Database)[9,10]. MeCP2 is a transcriptional regulatory protein, and in its absence, a large number of genes exhibit abnormal expression with implications in the balance between synaptic excitation and inhibition[11].

Although the majority of RTT patients have mutations in the *MECP2* gene[10], approximately 5% of classical RTT and 25% of variant RTT patients are negative for *MECP2* mutation[6,12]. In this variant RTT group of patients, some have mutations in other genes that are also associated with RTT: cyclin-dependent kinase-like5 (*CDKL5*; MIM *300203), which is described in individuals with an early seizure onset variant of RTT[13], and Forkhead box protein G1 (*FOXG1*; MIM *164874), which is responsible for the congenital variant of RTT[14]. In addition, with

[1]Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain. [2]Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain. [3]Centro Nacional de Análisis Genómica (CNAG-CRG), Center for Genomic Regulation, Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [4]Neurology Service, Hospital Sant Joan de Déu, Barcelona, Spain. [5]Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain. [6]CIBER-ER (Biomedical Network Research Center for Rare Diseases), Instituto de Salud Carlos III, Madrid, Spain. *A comprehensive list of consortium members appears at the end of the paper. Correspondence and requests for materials should be addressed to J.A. (email: jarmstrong@sjdhospitalbarcelona.org)

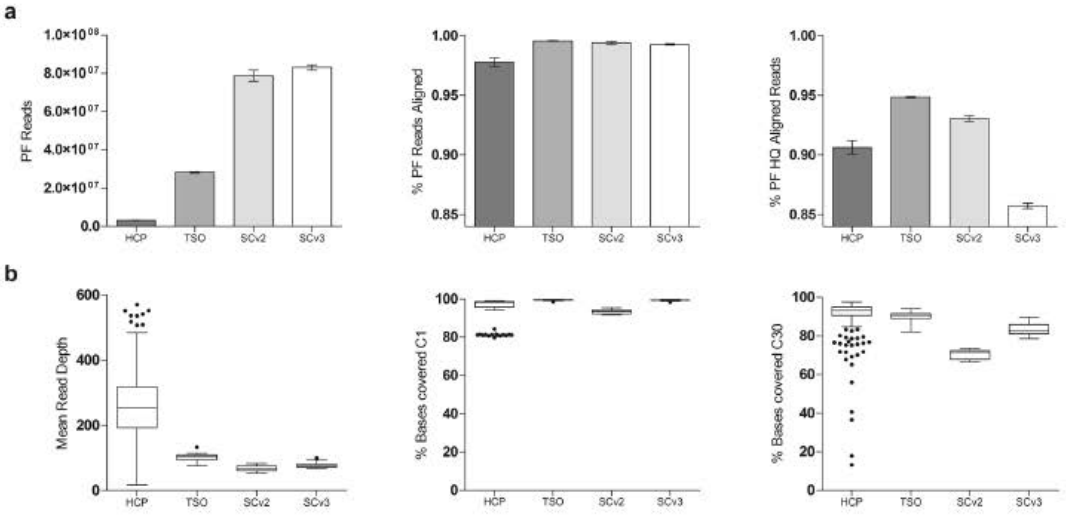


Figure 1. Comparison of main coverage metrics. Average of all samples analyzed for the four different approaches. (**a**) Bar plots with 95% confidence interval for the four approaches. Alignment metrics: passing filter (PF) reads; percentage of PF reads that aligned to the reference sequence; and percentage of PF reads that were aligned to the reference sequence with a mapping quality of Q20 or higher signifying. (**b**) Tukey boxplots: Mean read depth; percentage of bases covered at C1; and percentage of bases covered at C30.

the introduction of next generation sequencing (NGS), other genes without previous relation to RTT have been associated with RTT-like phenotypes, such as myocyte-specific enhancer factor 2C (*MEF2C*; MIM *600662) and transcription factor 4 (*TCF4*; *602272)[15]. However, the etiology of a subset of patients with a clinical diagnosis of RTT or RTT-like symptoms remains unknown.

In recent years, NGS has emerged as a potentially powerful tool for the study of this type of genetic disease[16,17]. Now, multiple genes can be sequenced at the same time and at a comparable cost to the Sanger analysis of only one single gene[18,19]. However, for genetic diagnostics, Sanger sequencing still remains necessary to validate the detected variants to avoid false positives.

Here, we present the retrospective results from our group using Sanger sequencing and three different approaches for library construction of NGS libraries. Our main goal was to assess the relative advantages and disadvantages of each methodology for diagnostic purposes.

Results

A total of 1577 patients with RTT-like syndrome were genetically analyzed between 1999 and 2016 at Sant Joan de Déu Hospital, Barcelona (see Fig. S1). These patients had been diagnosed following the usual clinical parameters[6] and according to the recently revised RTT Search International Consortium criteria and nomenclature[6]. Throughout the Sanger period, 84% of RTT classical phenotypes and 16% of atypical phenotypes were recruited for study. From 2012, a wider group of patients was recruited, not only RTT classical and atypical but also with RTT-like features, due to the incorporation of the NGS technologies.

Four approaches were used to genetically analyze all RTT patients who were recruited. A total of 1341 patients were studied based on Sanger sequencing (SS). The high throughput approaches were based on three distinct gene library preparations. First, 242 patients were studied using the Haloplex Custom Panel (HCP, Agilent Technologies) (Santa Clara, California), including 46 that came from a negative study of SS. Second, 51 patients were studied using the TruSight One panel (TSO, Illumina) (San Diego, California), including 11 who came from a negative study of HCP. Finally, 22 patients declared as negative-SS and 3 declared as negative-HCP were studied using the SeqCap EZ Human Exome v2.0 and v3.0 (WES, Roche NimbleGen) (Madison, Wisconsin). SS was used to study the exons and surroundings of the three genes associated with RTT (*MECP2*, *CDKL5* and *FOXG1*). HCP is a custom panel of 17 genes designed to cover the exons and surroundings of genes associated with RTT-like phenotypes. TSO targets 4,813 genes associated with a clinical phenotype, including RTT-like phenotypes. The WES kit was used to target all human coding exons, including genes covered by panels.

Run and Mapping Quality. A summary of quality control data is represented in Fig. 1a. The total number of passing filter reads (PF reads) was approximately 9 times higher in TSO (approx. 28 million reads) than HCP (approx. 3 million reads), whereas WES (approx. 80 million reads) was 3 times higher than TSO. The percentage of unique–mapping reads aligned to the reference sequence was higher in TSO (99.5%) and WES (99.4% and 99.2%) than HCP (97.8%). In addition, the unique–mapping reads aligned with a mapping quality Q20 or higher, indicating that the aligner estimates a 1/100 (or smaller) chance of a wrong alignment, was also higher in TSO (94.8%) and WES (93.0 and 85.7%) than HCP (90.6%).

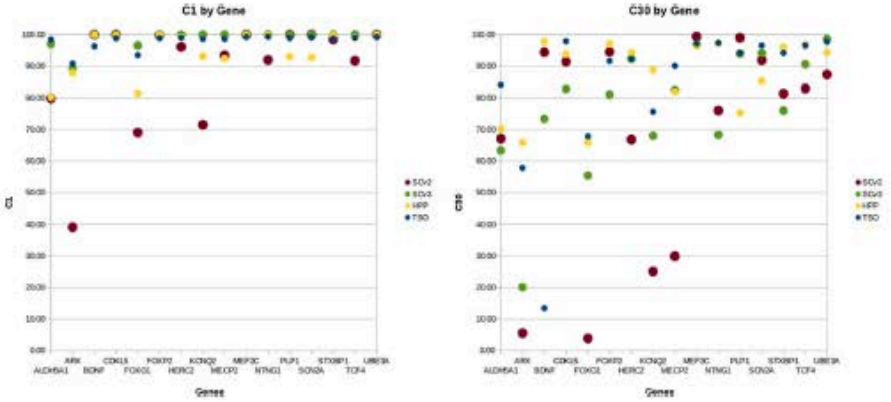


Figure 2. Comparison of coverage in 17 RTT-like genes. Scatter plots of average of the coverage at C1 and C30 of all samples analyzed for the four different approaches.

Target Regions, Read Depth and Coverage. The diagnosis regions of interest for a given set of genes that are related to RTT-like features is defined as the sum of the targets defined by the coding bases of the exons plus a 25 bp flanking region. Three sets of genes of interest were defined where each set included the previous one: RTT list, which included the 3 genes related to RTT; RTT-like list, 3+60 genes related to RTT-like disease; and RTT+EEP list, 3+14 RTT-like genes and 526 genes related to RTT-like features and epileptic encephalopathy (EEP) (Table 1S). Depending on the NGS methods and analyses performed, we evaluated the set of 17 genes common to both panels (TSO and HCP) and the WES, or the set of 605 genes most included in the TSO and all of them in the WES. The gene lists are shown in Supplementary Table 1S.

The performance of the three approaches was compared as if they were four since the WES samples were captured with two different kits, i.e., with SureSelect version 2.0 in 2011 (SCv2) and version 3.0 (SCv3) in 2014. In this study, 30 reads per base was considered the minimum coverage (C30) for high-sensitivity heterozygote detection. For the RTT-like list, average mean read depths of 262×, 99×, 67× and 77× were obtained by HCP, TSO, SCv2 and SCv3, respectively. Ninety percent of the targeted bases were covered at C30 by HCP, and for TSO, 88% of the bases were, while WES was 70% in 2011 and 84% in 2014. These mean coverage results are further compared in Fig. 1b, displaying variation among samples with the four methods referring to the mean read depth and bases covering C1 and C30. Both the coverage and the uniformity of the capture sequencing were better in TSO and WES than HCP. Deepening the coverage data, we analyzed the C1 and C30 of the 16 genes related to an RTT-like phenotype included in HCP considering the four approaches. For the C1 and C30 plots (Fig. 2), we used the raw CCDS coding exon coordinates as common references to compare as fairly as possible the different capture approaches. In this analysis, we discarded the *SHANK3* gene since this gene does not have a CCDS ID (Fig. 2).

Variant detection. According to diagnostic quality standards, all regions not reaching the required 30× must be Sanger sequenced; from HCP samples, 2 regions of the *MECP2* gene were sequenced by SS. Non-targets were sequenced by SS in TSO and WES. To identify the potential mutations, we checked the variants by matching their affected phenotypes and inheritance patterns of respective genes checked by SS of the index cases and their progenitors. Moreover, we considered the pathogenicity predictors (Mutation Tester, Polyphen-2 and Sorts Intolerant From Tolerant) and reviewed the literature, the RettBASE: MECP2 Variation Database, the Exome Aggregation Consortium (ExAC), HGMD® Professional 2016.4 and The Database of Short Genetic Variation (dbSNP). The potentially pathogenic variants detected in the genes that were not *MECP2*, *CDKL5* or *FOXG1* are shown in Table 1. All results are shown in Supplementary Table S2.

Sanger Sequencing-SS. A total of 1341 patients were studied by SS between 1999 and 2012, and 375 (22 by MLPA) were genetically diagnosed (28%) with mutations in *MECP2*, *CDKL5* or *FOXG1*. During this period of time, the workflow was to study the four exons of the *MECP2* gene to detect SNVs or short indels and MLPA for gross rearrangements. Excluding large rearrangements detected by MLPA, a total of 293 patients with RTT classic and 36 with atypical phenotypes had mutations in *MECP2*. Then, the *CDKL5* and *FOXG1* genes were studied in the patients without mutations in *MECP2*. No patient diagnosed with RTT-classic had mutations in *CDKL5* or *FOXG1*, and 15 patients with atypical RTT had mutations in *CDKL5* and 9 in the *FOXG1* gene.

Haloplex Custom Panel - HCP. A total of 242 patients were studied with HCP between 2012 and 2016, and 53 (6 by MLPA) patients were genetically diagnosed (22%). Excluding large rearrangements detected by MLPA, 29 patients had mutations in RTT genes and 90% in *MECP2*. From 18 patients with mutations in RTT-like genes, it is remarkable that the majority of these mutations were in the *STXBP1* gene, which is associated with early infantile epileptic encephalopathy (EIEE4)[20], and the *TCF4* gene, which is associated with Pitt–Hopkins syndrome[21].

Num. Patients	Gene	OMIM number	Transcript	Type of seq. Change	cDNA change	Protein change	dbSNP	Mutation taster, SIFT, PROVEAN, PolyPhen-2 scores
Potentially pathogenic mutation detected by HCP								
2	STXBP1	602926	NM_003165	Missense	c.874C>T	p.Arg292Cys	—	Disease causing, 0, 0.996, −7.58
2	STXBP1	602926	NM_003165	Missense	c.875G>A	p.Arg292His	rs796053361	Disease causing, 0, 1, −4.74
1	KCNQ2	602235	NM_172107	Missense	c.593G>A	p.Arg198Gln	rs796052621	Disease causing, 0, 1, −3.58
1	KCNQ2	602235	NM_172107	Missense	c.637C>T	p.Arg213Trp	rs118192203	Disease causing, 0, 1, −7.19
1	SLC2A1	138140	NM_006516	Missense	c.805C>T	p.Arg269Cys	rs200247956	Disease causing, 0, 1, −7.79
1	STXBP1	602926	NM_003165	Missense	c.1216C>T	p.Arg406Cys	rs796053367	Disease causing, 0, 1, −7.86
1	STXBP1	602926	NM_003165	In-frame deletion	c.124_126delTCC	p.Ser42del	—	Disease causing, NA, NA, −10.71
1	STXBP1	602926	NM_003165	Splicing variant	c.326-3C>G	Miss-splicing	—	NA, NA, NA, NA
1	STXBP1	602926	NM_003165	Missense	c.704G>A	p.Arg235Gln	—	Disease causing, 0, 1, −3.79
1	TCF4	602272	NM_001243236	In-frame indel	c.1169_1175delTAGAAAGinsAAA	p.Leu390Ter	—	Disease causing, NA, NA, NA
1	TCF4	602272	NM_001243236	Missense	c.1733G>A	p.Arg578His	rs121909123	Disease causing, 0, 1, −4.73
1	TCF4	602272	NM_001243236	Nonsense	c.1774C>T	p.Gln592Ter	—	Disease causing, NA, NA, NA
1	TCF4	602272	NM_001243236	Frameshift deletion	c.514_517delAAAG	p.Lys172PhefsTer61	rs398123561	Disease causing, NA, NA, NA
Potentially pathogenic mutation detected by TSO								
1	MEF2C	600662	NM_001193347	Missense	c.48C>G	p.Asn16Lys	—	Disease causing, 0.013, 0.995, −5.35
1	MEF2C	600662	NM_001193347	Frameshift deletion	c.989_990delGT	p.Gly330AspfsTer7	—	Disease causing, NA, NA, NA
1	SCN2A	182390	NM_001040142	Missense	c.3631G>A	p.Glu1211Lys	rs387906684	Disease causing, 0, 0.995, −3.82
1	SCN2A	182390	NM_001040142	Missense	c.5317G>A	p.Ala1773Thr	—	Disease causing, 0, 1, −3.68
1	SYNGAP1	603384	NM_006772	Frameshift deletion	c.2019delA	p.Thr674ProfsTer36	—	Disease causing, NA, NA, NA
1	SYNGAP1	603384	NM_006772	Frameshift deletion	c.1782delC	p.Leu595CysfsTer55	rs587780470	Disease causing, NA, NA, NA
Potentially pathogenic mutation detected by WES								
1	CACNA1I	608230	NM_021096	Missense	c.4435C>T	p.Leu1479Phe	—	Disease causing, 0.397, 0.756, −1.53
1	CHRNA5	118505	NM_000745	Missense	c.748C>A	p.Pro250Thr	—	Disease causing, 0.301, 1, −6.07
1	GABBR2	607340	NM_005458	Missense	c.1699G>A	p.Ala567Thr	—	Disease causing, 0.002, 0.999, −3.48
1	GRIN2B	138252	NM_000834	Missense	c.1657C>A	p.Pro553Thr	—	Disease causing, 0.001, 0.975, −6.8
1	HCN1	602780	NM_021072	Missense	c.1159G>T	p.Ala387Ser	—	Disease causing, 0.002, 0.767, −2.76

Table 1. Potentially pathogenic and causative SNVs detected, excluding RTT genes. Variant effect predictors web tools used: Mutation taster (http://www.mutationtaster.org/); SIFT-PROVEAN, SIFT scores ranged from 0–1, where 0 is predicted to be most damaging and Protein Variation Effect Analyzer (PROVEAN) score ≤ -2.5, the protein variant is predicted to have a "deleterious" effect, while if the PROVEAN score is > -2.5, the variant is predicted to have a "neutral" effect (http://provean.jcvi.org/genome_submit_2.php); and Polyphen-2, ranged from 0–1, where 1 is most likely to be damaging (http://genetics.bwh.harvard.edu/pph2/).

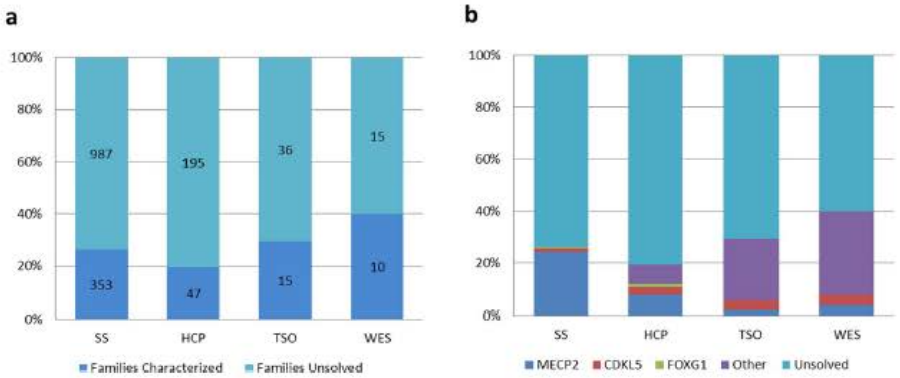


Figure 3. Statistics results for genetic diagnosis. Each columns represents different detection methodology used (SS, HCP, TSO and WES) (a) Percentage and total of families characterized and unsolved by the four different approaches. (b) Percentage of patients with *MECP2*, *CDKL5*, *FOXG1*, other genes with pathogenic mutations and the unsolved cases for the Sanger Sequencing (SS) and the three NGS approaches. Total of mutations found are detailed in Table S2.

TruSight One panel - TSO. Fifty-one patients were studied by TSO since 2015, and 15 patients were genetically diagnosed (29%). Three patients had mutations in RTT genes and 12 in RTT-like genes. In these 12 patients, we detected 2 SNVs in *MEF2C* and 2 in *SCN2A* genes, which are associated with mental retardation, stereotypic movements, epilepsy, and/or cerebral malformations[22–24]; 2 SNVs in the *SYNGAP1* gene, which is associated with mental retardation[24,25]; a deletion of *IQSEC2* and *KDM5C* genes, which are associated with mental retardation[26]; a gross deletion in chr15 (chr15:22,833,395-28,567,298), which is associated with Prader-Willi syndrome[27]; and a duplication in chr14 (14q32.11-q32.33(90949120-107287505)), which is associated with mental retardation and development delay[28]. All CNVs detected through the read depth and detected by NGS were confirmed by CGHarray.

***Whole Exome Sequencing* - WES.** Twenty-five patients were studied by WES in 2011 (SCv2) and 2014 (SCv3). Five occurred in genes previously associated with neurodevelopmental disorders with features similar to those of RTT syndrome. We identified four mutations in genes such as *HCN1*[29] and *GRIN2B*[30], which are associated with early infantile epileptic encephalopathy; *SLC6A1*, which is associated with epilepsy and myoclonic-atonic seizures[31]; *TCF4*, which is associated with Pitt–Hopkins syndrome[21]; and *SCN1A*, which is associated with Dravet syndrome[32]. The other potentially pathogenic variants that were identified occurred in genes that had not been linked to any genetic disorder[33]. However, there was an enrichment of genes with a potential role in neuronal biology and functionality, such as the gamma-aminobutyric type B receptor subunit 2 (*GABBR2*), the neuronal acetylcholine receptor subunit alpha-5 (*CHRNA5*), and the neuronal voltage-gated calcium channel (*CACNA1I*).

Genetic diagnosis. The total of cases and statistical results of genetic diagnoses of the RTT-like patients are listed in Fig. 3a. We identified the genetic cause for 353/1341 patients studied with SS. All characterized patients (26.3%) had mutations in RTT genes (*MECP2*, 24.5%; *CDKL5*, 1.1%; *FOXG1*, 0.7%). Forty-seven of 242 patients (19.4%) were genetically diagnosed by HCP; for these patients, 7.9% had mutations in the *MECP2* gene, 2.9% in *CDKL5*, 1.2% in *FOXG1* and 7.4% in other genes related to RTT-like phenotypes. Fifteen of 51 patients (29.4%) were genetically diagnosed by TSO; for these patients, 2.0% had mutations in the *MECP2* gene, 3.9% in *CDKL5* and 23.5% in other genes related to RTT-like phenotypes. Ten of 25 patients (40.0%) were genetically diagnosed by TSO; for these patients, 4.0% had mutations in the *MECP2* gene, 4.0% in *CDKL5* and 32.0% in other genes related to RTT-like phenotypes (Fig. 3b).

In addition of these results, it has been identified by MLPA twenty four patients with gross deletions which are not included in the statistical results. Twenty one patients with gross deletions in MECP2, where one have a deletion of exon 1 and 2, six patients in exon 3 and fourteen patients in exon 3 and 4. Two patients have a deletion in *CDKL5*, one with a deletion in exon 1 and 2 and other in exon 8. And three patients with a deletion FOXG1.

Discussion

Over the last few years, several genes have been associated with RTT-like phenotypes[17,33–37] due to the incorporation of NGS. Traditional detection methods for individual genes such as SS can only provide a limited mutation spectrum of the disease, consuming a great amount of time. NGS, which has revolutionized molecular genetics research, is a high-throughput method capable of rapidly sequencing a large number of genes in parallel and providing large datasets[33–38]. This study reports a comparison among different sequencing methods used across the last decade. We first developed a custom panel for molecular diagnosis of RTT and RTT-like phenotypes covering 17 genes related to RTT-like disease and then used a commercial panel that includes these genes and others also

associated with the phenotype. Furthermore, we performed WES to identify unsolved families via SS or a custom panel.

High performance standards are essential in a clinical diagnostic setting. Even though the mean read depth in HCP is ostensibly higher than other methods, with regions covering more than 500×, the uniformity of the capture is uneven. Figure 1b shows that some samples from different captures are in the border or below the quality standards of diagnosis. In spite of the fact that TSO has a slightly lower coverage at C30 than HCP, the homogeneity and uniformity of all the samples analyzed is more accurate. Even though the number of target region bases not reaching C30 is much higher in WES than in panel sequencing, it should be noted that the quality of the analysis improved in these few years (2011 to 2014). Moreover, these coverage comparatives could be inequitable, and we have to take into account that these WES are not performed with very high coverage and we cannot eliminate duplicates of PCR of the HCP. Thus, we not only compared the C30 but also the C1, which is the region that could potentially be covered if adequate coverage was made. Comparing the C1, TSO and SCv3 could potentially be better coverage than HCP and SCv2[39].

A large cohort of Spanish patients who exhibited clinical features associated with RTT or RTT-like phenotypes was recruited over three decades. Before the implementation of NGS, the cohort of patients was less clinically homogenous. Today, the cohort presents an increase in the number of patients who were RTT-like compared to classic or atypical RTT. This is likely due to the limitation of the technique because it takes time to analyze the 3 genes associated with RTT using SS and clinicians only send samples from patients who have fulfilled the clinical criteria as RTT.

Regarding additional information of this study, with a more extensive panel, the rate of the families characterized was higher. We characterized 40% of the families analyzed by WES and only 26% by SS. It is obviously less costly to sequence a panel of patients than an exome per trio (patient and progenitors) due to the much smaller capture region. Moreover, the development of panels targeting most clinically relevant genes, such as TSO or HCP, would be more cost-effective than standard WES. However, it must be taken into consideration that our HCP panel required a cost for development and validation, and TSO and WES are both labelled for research use only. For these reasons, proper validation by the laboratory is required before implementation in patient diagnoses. On the other hand, data obtained in a HCP or TSO are easier to interpret than WES results and also the computational requirements for a panel of a few genes is lower than a WES[40].

Regarding the NGS results, the *MECP2* gene remains the major mutated gene in our cohort (82% of all positive results), followed by *CDKL5* (6%) and *FOXG1* (3%) genes. One of the important findings is that two genes related to other phenotypes were found to be more frequently mutated than in other RTT-like phenotypes. Eight patients had pathogenic mutations in the *STXBP1* gene. The clinical features in patients with a mutation in this gene, such as developmental regression in the neonatal period or infancy, hypotonia, poor visual pursuit, seizures and epileptic encephalopathy, could fit perfectly with the RTT phenotype. Five patients had pathogenic mutations in the *TCF4* gene and one in the *UBE3A* gene, which are associated with Pitt-Hopkins syndrome and Angelman syndrome, respectively. Mutations in these genes cause severe neurologic features such as poor or absent speech development, delayed motor development, seizures, and hypotonia with an onset during the first year of life. Moreover, the patients with Pitt-Hopkins have morphologic characteristics, such as deep-set eyes and fleshy ears, but in our cohort, we found patients without any of these features. We also found 2 patients with mutations in the *MEF2C* gene (mental retardation, stereotypic movements, epilepsy, and/or cerebral malformations) and 2 patients with mutations in the *SYNGAP1* gene (mental retardation, autosomal dominant 5). The fact that there are the same neurologic features commented on previously highlights that *SYNGAP1* has behavioral psychiatric manifestations identical to RTT, that is, autism features and regression of motor development.

It is also remarkable that, in addition to *STXBP1* (EEIE4), we found mutations in genes related to EEIE: *KCNQ2* (EEIE7), *SCN1A* (EEIE6), *SCN2A* (EEIE11), *GRIN2B* (EEIE27) and *HCN1* (EEIE24) genes. Furthermore, we also found members of the solute carrier families related to epilepsy: *SLC6A1* (myoclonic-atonic epilepsy) and *SLC2A1* (idiopathic generalized epilepsy) genes[33,35].

Finally, mutations in genes without related phenotypes were detected by WES: *GABBR2, CHRNA5* and *CACNA1I*. Even though these genes are quite unknown, they have a potential role in neuronal biology and functionality. Further functional studies are required to consider these as the genetic cause of the disease. The clinical characteristics and detected variants of these patients are summarized in Lucariello *et al.*[33].

In this study, we focused on the comparison of a single nucleotide variant (SNV) and insertion/deletion detection, since these approaches are highly sensitive and specific to detect these types of mutations. Therefore, our diagnostics workflow includes MLPA (Multiplex Ligation-dependent Probe Amplification) of the *MECP2*, *CDKL5*, *FOXG1* and *TCF4* genes. aCGH (array comparative genomic hybridization) can be used for patients who have to be analyzed by WES. Although copy number variations (CNV) are difficult to detect with these NGS methods, we performed a preliminary study with TSO read depths of patients without detecting SNVs, trying to detect CNVs via bioinformatics methods. We were able to detect 3 *de novo* pathogenic CNVs corroborated later by aCGH: two deletions, one of the *IQSEC2* gene and one of the *KDM5C* gene, and another in chr15:22,833,395-28,567,298. One gross duplication in the long arm of chr14:90,949,120-107,287,505 was found.

In summary, the genetic study by NGS allows study of a larger number of genes associated with RTT-like symptoms simultaneously, allowing genetic study of a wider group of patients[10,41]. These detected variants identified by NGS may modify the initial clinical diagnosis to other neurodevelopmental syndromes, or determine new candidate genes related to RTT-like symptoms, providing the clinician with more information and clues that could help in the prevention of future symptoms or in the pharmacologic therapy. For instance, the use of D-serine as a dietary supplement for the enhancement of glutamatergic neurotransmission and/or excitatory/inhibitory neurotransmitter imbalance affected in patients with mutations in *N*-Methyl-D-aspartate receptors such as *GRIN2B* gene[42]. We could conclude that TSO has the best cost-efficiency of all technologies used and could offer timely responses for clinical diagnosis. However, performing a WES in families to characterize the

family and identify new candidate genes should be included in the target panels. In addition, study of the progenitors remains essential for their characterization as well as the need for functional studies in newly discovered genetic variants.

Material and Methods

Patients and DNA extraction. A cohort of 1577 Spanish patients who exhibited clinical features associated with RTT or RTT-like phenotypes was recruited at Sant Joan de Déu Hospital in Barcelona, Spain, from different Spanish Hospitals. Genomic DNA samples were extracted from peripheral blood leukocytes using the Puregene DNA Isolation kit (Gentra System, Minneapolis, USA) following the manufacturer's instructions. All DNA samples were quantified using a Qubit 2.0 Fluorometer (Invitrogen), and the DNA purity was quantified by calculating the absorbance ratio (*A*260/280) with a NanoDrop 1000 (Thermo Fisher).

Ethical issues. The study was approved by the ethical committees of Hospital Sant Joan de Déu, CEIC: Comité d'Ètica d'Investigación Clinica- Fundació Sant Joan de Déu (internal code: PIC-101-15). Patients or their parents signed informed consent for genetic studies, and tissue samples from patients and controls were obtained according to the Helsinki Declaration of 1964, as revised in 2001.

Molecular analysis. A total of 1341 patients were sequenced by SS for the 3 RTT genes (*MECP2*, *CDKL5* and *FOXG1*); 242 patients by HCP; 51 patients by TSO; and 24 patients with their healthy progenitors by WES. All patients analyzed by WES had an aCGH performed first with a normal or inconclusive profile.

Copy Number Variations - CNVs. *Multiplex Ligation-dependent Probe Amplification (MLPA).* *MECP2*-MLPA was performed using a SALSA kit P015, *CDKL5*-MLPA with a SALSA kit P189 and *FOXG1-TCF4*-MLPA with a SALSA kit P075 (MRC-Holland, Amsterdam, The Netherlands) according to the manufacturer's instructions.

Array Comparative Genomic Hybridization (aCGH). The aCGH analysis was performed at Bioarray (Genetic diagnosis, Alicante, Spain) using two different platforms: human genome Cytoarray Plus 180K and 400K (Agilent Technologies, Santa Clara, CA, USA). All genomic coordinates are in build GRCh37/hg19.

NGS: Library Preparation. Library preparation was conducted according to the manufacturers' instructions. HCP and TSO panels were created at Sant Joan de Déu Hospital and WES at the National Center for Genomic Analysis (CNAG).

Haloplex Custom Panel (HCP). We designed a custom-made panel with 17 genes associated with a RTT-like phenotype based on the evidence curated in the Online Mendelian Inheritance in Man (OMIM). The genes are shown in Supplementary Table S3. Amplicon libraries were prepared using the Agilent HaloPlex Target enrichment system, for Illumina paired-end multiplexed sequencing platforms (Agilent Technologies), according to the manufacturer's sample preparation protocol. Briefly, 225 ng of genomic DNA was digested with restriction enzymes. The hybridization was performed for 3 hours at 54 °C, and the circularized target DNA-HaloPlex probe hybrids, containing biotin, were captured on streptavidin beads (HaloPlex Magnetics Beads, Agilent Technologies). The DNA with adaptor-modified ends was PCR amplified (number of cycles depended on the lot, Herculase II fusion DNA polymerase, Agilent). The amplified target DNA was purified using AMPure XP beads (Beckman Coulter Genomics, GENEWIZ, New Jersey, USA). All of the DNA samples were individually indexed. Amplification of the libraries was performed on a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems). The restriction digestion and amplicon library quantities were quality evaluated using a Bioanalyzer High Sensitivity DNA Assay kit in an Agilent 2100 Bioanalyzer (Agilent Technologies) and quantified using a Qubit 2.0 Fluorometer (Invitrogen).

TruSight One Panel (TSO). TruSight One Sequencing Panel (Illumina, San Diego, CA) targeted 4,813 genes associated with a clinical phenotype. Libraries were generated using the TruSight One™ Sequencing Panel kit with the TruSight One™ Sequencing Panel (Illumina), according to the manufacturer's sample preparation protocol. Briefly, 50 ng of each DNA sample was enzymatically fragmented and adapter sequences were added to the ends. The fragmented DNA was purified, and barcodes and common adapters required for cluster generation and sequencing were PCR-added. After cleanup, 500 ng of each of the 12 DNA libraries was pooled. Then the libraries were hybridized twice to specific capture probes; the unhybridized material was washed away, and the captured fragments were amplified using PCR followed by purification. The enriched libraries were quantified using a Qubit 2.0 Fluorometer, and their quality was evaluated using a Bioanalyzer 2100 and the High Sensitivity DNA Kit (Agilent Technologies). Libraries were diluted and pooled to obtain the final sequencing equimolar pool.

Whole Exome Sequencing (WES). For whole exomes from 2011, the sample preparation for capturing approximately 44 Mb of selected human genome regions was performed according to the NimbleGen SeqCap EZ Exome Libray SR protocol, v2.2, for Illumina paired-end sample libraries with modifications included in draft v1.4 (February 2011). In brief, 1.0 µg of genomic DNA was sheared on a Covaris™ E210 instrument (Covaris). The fragment size (150–400 bp) and the quantity were confirmed with the Agilent 2100 Bioanalyzer 1000 chip (Agilent). Fragmented DNA was prepared using an Illumina TruSeq DNA Sample Preparation Kit (Illumina) following the protocol described in the Illumina TruSeq DNA Sample Preparation Guide (revision A, November 2010) with the exception of the fragment size selection using an Agencourt SPRI XP Kit (Beckman Coulter). Instead of the standard Illumina PCR enrichment step, amplification via pre-capture LM-PCR (8 cycles) was performed with a SeqCap EZ Human Exome Kit v2.0 (Roche NimbleGen). One microgram of the amplified

library was hybridized to EZ probes at 47 °C for 72 hrs. After washing and recovery of the captured DNA, the library was amplified through post-capture LM-PCR (18 cycles). The final product was quality controlled on a Bioanalyzer DNA 1000 chip, and the success of the enrichment was measured with a qPCR SYBR Green assay on a LightCycler® 480 Instrument (Roche), evaluating one genomic locus with pre- and post-captured material.

For whole exomes from 2014, the NimbleGen SeqCap EZ v3.0 system for exome enrichment was used and pre-capture multiplexing was applied following the manufacturer's protocol version 4.2. Briefly, 1 µg of genomic DNA was fragmented with Covaris ™ E210 and used for ligation of the adapters containing Illumina specific indexes with a KAPA Library Preparation kit (Kapa Biosystems). Adapter ligation DNA fragments were enriched by 7 cycles of pre-capture PCR using KAPA HiFi HotStart ReadyMix (2×) (Kapa Biosystems) and analyzed on an Agilent 2100 Bioanalyzer with the DNA 1000 assay. Five libraries were pooled with a combined mass of 1250 ng for the baits hybridization step (47 °C; 68 hrs). After washing (47 °C), the multiplexed captured library was recovered with capture beads and amplified with 14 cycles of post-capture PCR using KAPA HiFi HotStart ReadyMix (2×). The size, concentration and quality of the captured library were determined using an Agilent DNA 1000 chip. The success of the enrichment was measured using a qPCR SYBR Green assay on a Roche LightCycler® 480 Instrument evaluating one genomic locus with pre- and post-captured material.

NGS: Run. HPC and TSO libraries were sequenced on an Illumina Miseq instrument and NextSeq 500, respectively, following the manufacturer's protocol, with a paired end run. Each WES (2011 and 2014) library was sequenced on an Illumina HiSeq 2000 instrument in a fraction of a sequencing lane following the manufacturer's protocol, with a paired end run of 2× 101 bp. Image analysis, base calling and quality scoring of the run were processed using the manufacturer's software Real Time Analysis (RTA 1.10.36 and RTA 1.13.48) and followed by generation of FASTQ sequence files by CASAVA.

NGS: Data Analysis. HPC and TSO data analyses were performed at Genycell Biotech S.L. (Madrid, Spain). Briefly, both panel sequencing reads were trimmed from the 3′ end up to the first base with a Phred quality >9 and were mapped to the Human genome reference v37 with decoy sequences [hs37d5] (Broad, ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/phase2_reference_assembly_sequence/hs37d5.fa.gz) using the BWA-MEM version 0.7.5a. (bio-bwa.sourceforge.net/). To calculate read statistics, we used Prinseq-lite version 0.20.3 lite (prinseq.sourceforge.net/) and Picard Calculate HS Metrics version 1.119 (http://picard.sourceforge.net) to calculate metrics. Alignment (.bam) files containing only properly paired and uniquely mapped reads were processed with Picard Mark Duplicates version 1.119 to remove duplicates, and local realignment was performed with Freebayes version 9.9.13 (https://github.com/ekg/freebayes)[43]. SAMtools version 0.1.19 (http://samtools.sourceforge.net/)[44] was used on the processed BAM files to call single nucleotide variants (SNVs) and small insertion deletions (INDELs). Functional annotations from Ensembl release 75 (GRCh37.75 database)[45] were added to the resulting.vcf files using the Genome Analysis Tool Kit (GATK) version 2.4 (https://www.broadinstitute.org/gatk/)[46] to annotate variants with dbSNP version 137 and Ensembl Variant Effect Predictor version 72 (http://www.ensembl.org/info/docs/tools/vep/index.html) to annotate the variants.

For each WES (2011 and 2014), sequencing reads were trimmed from the 3′ end up to the first base with a Phred quality >9 and were mapped to the Human genome reference v37 with decoy sequences [hs37d5] using the GEM toolkit[47]. Alignment (.bam) files containing only properly paired and uniquely mapped reads were processed with Picard tools version 1.110 (http://picard.sourceforge.net) to remove duplicates, and local realignment was performed with the Genome Analysis Tool Kit (GATK) version 3.1 (https://www.broadinstitute.org/gatk/)[46]. For alignments and coverage metrics, it has been remapped with BWA-MEM version 0.7.5a[48]. To call single nucleotide variants (SNVs) and small insertion deletions (INDELs) on the processed alignment (.bam) file, we used SAMtools version 0.1.19 (http://samtools.sourceforge.net/)[44]. Functional annotations from Ensembl release 75 were added to the resulting VCF using snpEff (http://snpeff.sourceforge.net/)[49]. SNPSift (http://snpeff.sourceforge.net/SnpSift.html)[58] was used to add information from dbSNP version 137, population frequencies from 1000 Genomes and the Exome Variant Server[51], the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS/) and a variety of conservation and deleteriousness predictions included in dbNSFP version 2.5 (http://sites.google.com/site/jpopgen/dbNSFP)[52].

Sanger validation. The identified variants and familial segregation studies were validated by SS. The specific primers were designed online by Primer3 version 0.4.0 (http://bioinfo.ut.ee/primer3-0.4.0/primer3/). The PCR products were sequenced using a Big-Dye® Terminator version 3.1 Cycle Sequencing Kit in an Applied Biosystems 3730/DNA Analyzer (Applied BioSystems, Waltham, Massachusetts, USA). The raw data were analyzed with Chromas trace viewer (http://technelysium.com.au/wp/chromas/). The primers used for SS are shown in Supplementary Table S4.

References

1. Rett, A. On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wien. Med. Wochenschr.* **116**, 723–726 (in German) (1966).
2. Hagberg, B., Aicardi, J., Dias, K. & Ramos, O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* **14**, 471–479 (1983).
3. Laurvick, C. L. *et al.* Rett syndrome in Australia: a review of the epidemiology. *J Pediatr* **148**, 347–352 (2006).
4. Diagnostic criteria for Rett syndrome. The Rett Syndrome Diagnostic Criteria Work Group. *Ann Neurol* **23**, 425–428 (1988).
5. Hagberg, B., Hanefeld, F., Percy, A. & Skjeldal, O. An update on clinically applicable diagnostic criteria in Rett syndrome. Comments to Rett Syndrome Clinical Criteria Consensus Panel Satellite to European Paediatric Neurology Society Meeting, Baden Baden, Germany, 11 September 2001. *Eur J Paediatr Neurol* **6**, 293–297 (2002).
6. Neul, J. L. *et al.* Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol* **68**, 944–950 (2010).

7. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**, 185–188 (1999).
8. Miltenberger-Miltenyi, G. & Laccone, L. Mutations and Polymorphisms in the Human Methyl CpG-Binding Protein MECP2. *Mut Genet* **135**, 1343–1354 (2003).
9. Christodoulou, J., Grimm, A., Maher, T. & Bennetts, B. RettBASE: the IRSA *MECP2* variation database — a new mutation database in evolution. *Hum Mutat* **21**, 466–472 (2003).
10. Neul, J. L. *et al.* Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* **70**, 1313–1321 (2008).
11. Chao, H. T. *et al.* Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* **468**, 263–269 (2010).
12. Neul, J. L. *et al.* Developmental delay in Rett syndrome: data from the natural history study. *J Neurodev Disord* **6**, 20 (2014).
13. Kalscheuer, V. M. *et al.* Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nat Genet* **35**, 313–315 (2003).
14. Ariani, F. *et al.* FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet* **83**, 89–93 (2008).
15. Armani, R. *et al.* Transcription factor 4 and myocyte enhancer factor 2C mutations are not common causes of Rett syndrome. *Am J Med Genet A* **158A**, 713–719 (2012).
16. Zhu, X. *et al.* Whole-exome sequencing in undiagnosed genetic disorders: interpreting 119 trios. *Genet Med* **17**, 774–781 (2015).
17. Sajan, S. A. *et al.* Enrichment of mutations in chromatin regulators in people with Rett Syndrome lacking mutations in MECP2. *Genet Med* **19**, 13–19 (2017).
18. Sikkema-Raddatz, B. *et al.* Targeted next-generation sequencing can replace Sanger sequencing in clinical diagnostics. *Hum Mutat* **34**, 1035–1042 (2013).
19. Kurian, A. W. *et al.* Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol* **32**, 2001–2009 (2009).
20. Carvill, G. L. *et al.* GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. *Neurology* **14**, 1245–1253 (2014).
21. Sweatt, J. D. Pitt-Hopkins Syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. *Exp Mol Med* **45**, e21 (2013).
22. Zweier, M. *et al.* Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum Mutat* **6**, 722–733 (2010).
23. Bienvenu, T., Diebold, B., Chelly, J. & Isidor, B. Refining the phenotype associated with MEF2C point mutations. *Neurogenetics* **14**, 71–75 (2013).
24. Carvill, G. L. *et al.* Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet* **45**, 825–830 (2013).
25. Mignot, C. *et al.* Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet* **53**, 511–522 (2016).
26. Fieremans, N. *et al.* Microdeletion of the escape genes KDM5C and IQSEC2 in a girl with severe intellectual disability and autistic features. *Eur J Med Genet* **58**, 324–327 (2015).
27. Buiting, K. Prader-Willi syndrome and Angelman syndrome. *Am J Med Genet C Semin Med Genet* **154C**, 365–376 (2010).
28. Thiel, C. T. *et al.* A de novo 7.6 Mb tandem duplication of 14q32.2-qter associated with primordial short stature with neurosecretory growth hormone dysfunction, distinct facial anomalies and mild developmental delay. *Eur J Med Genet* **51**, 362–367 (2008).
29. Nava, C. *et al.* De novo mutations in HCN1 cause early infantile epileptic encephalopathy. *Nat Genet* **46**, 640–645 (2014).
30. Lemke, J. R. *et al.* GRIN2B mutations in West syndrome and intellectual disability with focal epilepsy. *Ann Neurol* **75**, 147–154 (2014).
31. Carvill, G. L. *et al.* Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. *Am J Hum Genet* **96**, 808–815 (2015).
32. Brunklaus, A. & Zuberi, S. M. Dravet syndrome–from epileptic encephalopathy to channelopathy. *Epilepsia* **55**, 979–984 (2014).
33. Lucariello, M. *et al.* Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype. *Hum Genet* **135**, 1343–1354 (2016).
34. Olson, H. E. Mutations in Epilepsy and Intellectual Disability Genes in Patients with Features of Rett Syndrome. *Am J Med Genet* **167A**, 2017–2025 (2015).
35. Gold, W. A. & Christodoulou, J. The utility of next-generation sequencing in gene discovery for mutation-negative patients with Rett syndrome. *Front Cell Neurosci* **9**, 266 (2015).
36. Lopes, F. *et al.* Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet 2016.* **53**, 190–199 (2016).
37. Marx, V. Next-generation sequencing: The genome jigsaw. *Nature* **501**, 263–268 (2013).
38. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491–498 (2011).
39. Samorodnitsky, E. *et al.* Comparison of Custom Capture for Targeted Next-Generation DNA Sequencing. *J Mol Diagn* **17**, 64–75 (2015).
40. Sun, Y. *et al.* Next-Generation Diagnostics: Gene Panel, Exome, or Whole Genome? *Hum Mutat* **36**, 648–655 (2015).
41. Bebbington, A. *et al.* Investigating genotype-phenotype relationships in Rett syndrome using an international data set. *Neurology* **70**, 868–875 (2008).
42. Soto, D. *et al.* Rett-like severe encephalopathy caused by a de novo GRIN2Bmutation is attenuated by D-serine dietary supplement. *Biol Psychiatry* **S0006-3223**, 31671–2, doi:https://doi.org/10.1016/j.biopsych.2017.05.028 (2017).
43. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907 [q-bio.GN]* (2012).
44. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (Oxford, England) **25**, 2078–2079 (2009).
45. Cunningham, F. *et al.* Ensembl 2015. *Nucleic Acids Res* **43**, D662–669 (2015).
46. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297–1303 (2010).
47. Marco-Sola, S., Sammeth, M., Guigó, R. & Ribeca, P. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat Methods* **9**, 1185–1188 (2012).
48. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997 [q-bio.GN]* (2013).
49. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
50. Cingolani, P. *et al.* Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front genet* **3**, 35 (2012).
51. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308–311 (2001).
52. Liu, X., Jian, X. & Boerwinkle, E. dbNSFP v2.0: a database of human non-synonymous SNVs and their functional predictions and annotations. *Hum mutat* **34**, E2393–2402 (2013).

Acknowledgements

We thank all patients and their families who contribute to this study. The work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01013); Crowdfunding program PRECIPITA, from the Spanish Ministry of Health (Fundación Española para la Ciencia y la Tecnología); Catalan Association for Rett Syndrome; Fondobiorett and Mi Princesa Rett. The molecular analyses of eleven WES in 2014 were funded by the CNAG's 2013 call "300 exomes to elucidate rare diseases". AGC is supported by the ISCIII grant: FIS PI15/01082.

Author Contributions

J.A. and M.P. conceived and supervised the study. S.V., P.P., N.B., E.G. and L.B., performed the experiments and collected the data. J.A., S.V., N.B., P.P., E.G., S.D., J.T. analyzed the results. J.A., M.'O.C., A.G.C., M.P. provided samples and patients' clinical and genetic information. S.V., J.A., J.T., S.D. wrote the manuscript. All the authors reviewed the article critically for intellectual content.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-11620-3

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017

Capítulo 2

Capítulo 2: Caracterización clínica y molecular de pacientes con grandes deleciones en *MECP2*

Aproximadamente el 5% de las mutaciones detectadas en *MECP2* son grandes reordenamientos que comprenden desde la deleción de un único exón hasta el gen completo. Para la mejora del diagnóstico y del pronóstico de las pacientes con grandes deleciones, hemos analizado los mecanismos que pueden llegar a producir estos grandes reordenamientos y hemos intentado establecer una correlación entre los distintos tamaños de las deleciones y la clínica que desarrollan las pacientes.

Publicación 3

Título: *Characterization of large deletions of the MECP2 gene in Rett syndrome patients by gene dosage analysis.*

Autores: Silvia Vidal, Ainhoa Pascual-Alonso, Marc Rabaza-Gairí, Edgar Gerotina, Núria Brandi, Paola Pacheco, Clara Xiol, Mercè Pineda, *Rett Working Group* and Judith Armstrong.

Referencia: *Molecular Genetics and Genomic Medicine*. 2019 Aug; 7(8): e793.

Resumen:

En este estudio hemos caracterizado a nivel molecular las deleciones detectadas por MLPA en el gen *MECP2* en 21 pacientes con RTT. Los puntos de rotura de estas grandes deleciones se acotaron y definieron mediante qPCR hasta que fue posible la amplificación del alelo delecionado mediante PCR larga. Esta metodología nos ha permitido caracterizar deleciones que van desde 1,235pb a 85kb, confirmando la eliminación parcial o total del gen *MECP2* en todas estas pacientes. Además, nuestros casos respaldan las evidencias que afirman que la mayoría de estos puntos de rotura ocurren en algunas regiones concretas del gen, en una pequeña región localizada en el exón 4 y una zona repetitiva en el intrón del gen. Estos datos moleculares junto con la información clínica nos han permitido proponer una correlación genotipo-fenotipo, importante para mejorar el asesoramiento genético proporcionado.

Received: 24 August 2018 | Revised: 24 January 2019 | Accepted: 16 May 2019

DOI: 10.1002/mgg3.793

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

Characterization of large deletions of the *MECP2* gene in Rett syndrome patients by gene dosage analysis

Silvia Vidal[1,2] | **Ainhoa Pascual-Alonso**[1,2] | **Marc Rabaza-Gairí**[1,2] | **Edgar Gerotina**[1,2] | **Nuria Brandi**[3] | **Paola Pacheco**[4] | **Clara Xiol**[1,2] | **Mercè Pineda**[1] | **Rett Working Group** | **Judith Armstrong**[2,4,5]

[1]Sant Joan de Déu Research Foundation, Barcelona, Spain

[2]Sant Joan de Déu Research Institute (IRSJD), Hospital Sant Joan de Déu, Esplugues de Lobregat (Barcelona), Spain

[3]Facultat de Medicina, Universitat de Barcelona, Barcelona, Spain

[4]Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain

[5]CIBER-ER (Biomedical Network Research Center for Rare Diseases), Instituto de Salud Carlos III, Madrid, Spain

Correspondence
Judith Armstrong, Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Pg. Sant Joan de Déu 2, planta 0, 08950 Esplugues de Llobregat, Barcelona.
Email: jarmstrong@sjdhospitalbarcelona.org

Funding information
This work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01159), Catalan Association for Rett Syndrome, Fondobiorett and Mi Princesa Rett.

Abstract

Background: Rett syndrome (RTT) is a developmental disorder with an early onset and X-linked dominant inheritance pattern. It is first recognized in infancy and is seen almost always in girls, but it may be seen in boys on rare occasions. Typical RTT is caused by de novo mutations of the gene *MECP2* (OMIM*300005), and atypical forms of RTT can be caused by mutations of the *CDKL5* (OMIM*300203) and *FOXG1* (OMIM*164874) genes.

Methods: Approximately 5% of the mutations detected in *MECP2* are large rearrangements that range from exons to the entire gene. Here, we have characterized the deletions detected by multiplex ligation-dependent probe amplification (MLPA) in the gene *MECP2* of 21 RTT patients. Breakpoints were delineated by DNA-qPCR until the amplification of the deleted allele by long-PCR was possible.

Results: This methodology enabled us to characterize deletions ranging from 1,235 bp to 85 kb, confirming the partial or total deletion of the *MECP2* gene in all these patients. Additionally, our cases support the evidence claiming that most of these breakpoints occur in some restricted regions of the *MECP2* gene.

Conclusion: These molecular data together with the clinical information enable us to propose a genotype–phenotype correlation, which is essential for providing genetic counseling.

KEYWORDS
large deletions, *MECP2*, Phenotype-genotype correlations, Rett syndrome

1 | INTRODUCTION

Rett syndrome (RTT; OMIM#312750) is a neurodevelopmental disorder with early onset that is most often found in girls. It is first recognized in infancy; a period of apparently normal development (up to the age of 6–18 months) is followed by a stagnation-regression characterized by a loss of purposeful hand use and speech, motor apraxia that may be associated with epilepsy and dysautonomic features, including disturbed breathing, sleep, and gastrointestinal motility(Hagberg, Aicardi, Dias, & Ramos, 1983). RTT has a worldwide incidence of 1:10,000 live female births and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

is the second leading cause of severe mental retardation in females.

Since 1999, numerous reports have supported the evidence that mutations in the *Methyl CpG binding protein 2* gene (*MECP2*; OMIM*300005) are the primary cause of classic RTT (Amir et al., 1999). MeCP2 is a transcriptional regulatory protein, and in its absence, a large number of genes exhibit abnormal expression with implications in the balance between synaptic excitation and inhibition (Kron et al., 2012).

The *MECP2* gene is localized in Xq28, contains four exons, and encodes two major functional domains namely: the methyl binding domain (MBD) (Nan, Meehan, & Bird, 1993) and the transcription repression domain (TRD), which contains a nuclear localization signal (NLS) (Singh, Saxena, Christodoulou, & Ravine, 2008). The *MECP2* translational initiation site was originally identified in exon 2, but a second translation initiation site was described in exon 1, which led to a new MeCP2 isoform (Mnatzakanian et al., 2004). MeCP2E1 is comprised of the exons 1, 3, and 4 while MeCP2E2 contains exons 2, 3, and 4; both forms comprise the MBD and TRD domains. MeCP2E1 is much more abundant in the brain while MeCP2E2 has a higher transcriptional expression level in the skeletal muscles, placenta, liver, and prostate gland (Liyanage & Rastegar, 2014).

No clear phenotype–genotype correlation has been identified in RTT patients (Bebbington et al., 2012; Neul et al., 2008; Scala et al., 2007). It has been reported that 95% of individuals affected by classic RTT have a loss of function in *MECP2*, but is less frequently seen in atypical RTT (Neul et al., 2010). There are eight common mutations of this gene that constitute approximately two-thirds of all mutations. Another small number of the patients carry a deletion ranging between 1 and 338 bp, in the C-terminal region (see RettBASE; http://mecp2.chw.edu.au/).

Soon after *MECP2* was identified as causative of RTT, several groups started to study the gene dosage of the cases that were negative for point mutations or small *indels* in the coding sequence of the gene. Thus, Southern Blot or MLPA followed by qPCR, long-PCR, and Sanger sequencing to narrow down the rearrangement, proved to be helpful in explaining approximately 10% of the mutations in those cases (Archer et al., 2006; Erlandson et al., 2003; Laccone et al., 2004; Ravn et al., 2005; Yaron et al., 2002). At our hospital, when taking into account all the cases diagnosed as RTT that have a mutation in *MECP2*, 4.5% of them have large rearrangements (Vidal et al., 2017), which is consistent with what has been reported in the literature (Hardwick et al., 2007).

Here, we present the molecular characterization of the breakpoints of the deletions detected in *MECP2* by MLPA in 21 RTT patients. The patients' clinical information was gathered as well, when available, in order to assess their severity with Pineda's score to determine a genotype–phenotype correlation and attempt to improve the genetic counseling for these and similar families.

2 | MATERIALS AND METHODS

2.1 | Patients and DNA samples

2.1.1 | Ethical compliance

Written informed consent was obtained from individuals legally responsible for the patients in accordance with appropriate ethics protocols for the analysis of genes related to RTT.

This study involved 21 patients clinically diagnosed with classic RTT who were negative for *MECP2* point mutations and small *indels* in the coding sequence. To evaluate the severity of the clinical presentation of each patient, a set of symptoms were measured using the clinical severity scores designed by Dr. Pineda (Monrós et al., 2001).

DNA was extracted from peripheral blood leukocytes using the Puregene DNA Isolation kit (Gentra System, Minneapolis, USA).

2.2 | MLPA analysis

All patients were analyzed by MLPA. MECP2-MLPA was performed with SALSA P015-D1, P015-E1 or P015-F1 kits (MRC-Holland, Amsterdam, The Netherlands) in accordance with the manufacturer's instructions. This assay covers all four *MECP2* exons and the flanking genes *IRAK1* (OMIM*300283), *L1CAM* (OMIM*308840), and *VAMP7* (OMIM*300053).

2.3 | Quantitative-PCR analysis (qPCR)

To narrow down the deletion breakpoints in each patient, we used real-time qPCR to test the relative copy number of various strategically designed amplicons located along the *MECP2* gene. Primers were designed from the genomic clone NM_004992.3 using Primer3 program (primer sequences and annealing sites in Supplementary Data S1). Briefly, our qPCR strategy was based on generating standard curves for each *MECP2* amplicon and for the autosomal reference gene *MTHFR* (OMIM*607093). These standard curves defined the relationship between the input DNA concentration and the C_t value.

The real-time qPCR was performed with the GoTaq Master Mix kit (Promega Corp., USA) for ABI 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA) and the PowerUp SYBR Green Master Mix kit for the QuantStudio 6 Flex Real-Time PCR System (both from Applied Biosystems, USA). All reactions were conducted in triplicate with the average of each triplicate group used for quantitative analysis. Product specificity was assessed by melting curve analysis.

TABLE 1 All the collected data from the patients; genomic information is based on the GRCh38/hg38 and the accession number NM_004992.3 (NG_007107.2) for *MECP2* gene

Patient ID	Clinical phenotype	Clinical severity score	Age of scoring, yr.	*MECP2* exons deleted	Breakpoint	Deletion Size	ICX
P1	Classic	8	11	4	c.1153_2387del	1,235 bp	72:28
P2	Classic	NA	NA	4	c.1157_3664del	2,508 bp	88:22
P3	Classic	11	10	4	c.1041_4447del	3,407 bp	64:36
P4	Classic	10[a]	17	4	c.1164_4665del	3,502 bp	87:13
P5	Classic	15	2	4	c.1164_1461+3282del	3,580 bp	58:42
P6	Classic	NA	NA	3 and 4	c.27-1125_1146del	3,001 bp	55:45
P7	Classic	NA	NA	3 and 4	c.27-5834_1166del	7,720 bp	58:42
P8	Classic	NA	NA	3 and 4	c.27-10677_1192del	12,599 bp	88:12
P9	Classic	NA	NA	3 and 4	c.(26+1_27-1)_(378_1385)del	≈ 11.5 kb	82:18
P10	Classic	17	10	3 and 4	c.27-7985_1209del	9,924 bp	51:49
P11	Classic	10	6	3 and 4	c.27-2950_1170del	4,850 bp	62:38
P12	Classic	13[a]	6	3 and 4	c.27-6312_1301delinsTG	8,343 bp	68:32
P13	Classic	14	4	3 and 4	c.(26+1_27-1)_(989_1241)del	≈19.15 kb	74:26
P14	Classic	NA	NA	3 and 4	c.27−15131_1461 + 7939del::insGGATCAGGT	25,261 bp	76:24
P15	Classic	13	8	3 and 4	c.(26+1_27-1)_(1461+4313_1461+5337)del	≈22.5 kb	83:17
P16	Classic	14	3	3, 4 and *IRAK1*	g.(154092184_154032557)_(154000172_153997133) del	≈ 50.7 kb	94:6
P17	Classic	10	10	3, 4 and *IRAK1*	g.(154092184_154032557)_(153986984_153969656) del	≈ 77.9 kb	88:12
P18	Classic	15	21	3, 4 and *IRAK1*	g.(154092184_154032557)_(154021813_154018990) del	≈ 28.3 kb	64:36
P19	Classic	15	16	3, 4 and *IRAK1*	c.[27-16409_1201delinsGGGGGCC; 1202_1460+2170inv; 1460+2171_1460+12766delinsTCTGCACGGGG]	18339bp and 10596bp	97:03
P20	Classic	14[a]	NA	4 and *IRAK1*	g.154030942_154003453del	27,589 bp	74:26
P21	Classic	17	NA	1 and 2	g.(154128954_154097731)_(154092184_154032557) del	≈ 85 kb	73:27

Abbreviation: NA, not available.
[a]Indicates the lack of information for some of the clinical features (See Supplementary Data S4); so the given score is the result of the data we were given by the clinicians.

The *MECP2* amplicon of interest and the *MTHFR* reference amplicon were amplified separately for each patient and for three normal female controls, yielding a copy number variant for each.

2.4 | Long-range PCR amplification and Sanger sequencing of deletion junctions

Once the deletions' breakpoints had been narrowed down to a sufficiently small region by qPCR, primer sites in the regions immediately flanking the breakpoints were selected for long-range PCR amplification. As the precise size of the junction fragment in each patient was unknown, several different PCR conditions were tested and optimized. Long-range PCR was performed with the Expand High Fidelity PCR System kit (Roche, Mannheim, Germany). This protocol was carried out in accordance with the manufacturer's instructions on a SimpliAmp Thermal cycler (Applied BioSystems, Waltham, MA). The PCR products were sequenced using a Big-Dye® Terminator version 3.1 Cycle Sequencing Kit in an Applied Biosystems 3,730/DNA Analyzer (Applied BioSystems, Waltham, MA). The raw data were analyzed with Chromas trace viewer (http://technelysium.com.au/wp/chromas/). The sequences of the junction fragments were aligned to the reference sequence of *MECP2* (NM_004992.3) using Genomatix diAlign® program (local multiple alignment; http://www.genomatix.de/cgi-bin/dialign/dialign.pl).

2.5 | X chromosome inactivation assay (XCI)

The XCI status of all 21 female patients was determined by the analysis of the methylation status of the highly polymorphic trinucleotide X-linked androgen receptor (*AR*; OMIM*313700) locus. For each subject, 50 ng of genomic DNA was digested separately with *Hpa*II restriction enzyme (New England Biolabs, Beverly, MA) in accordance with the manufacturer's instructions. A region between 252 and 327 bp of the locus was PCR amplified from digested and undigested DNA using fluorochrome-labeled primers. Samples were electrophoresed on an ABI Prism Genetic Analyzer 3130, and the peak areas were quantified using Gene Mapper v4.0 software (Applied Biosystems, Foster City, CA).

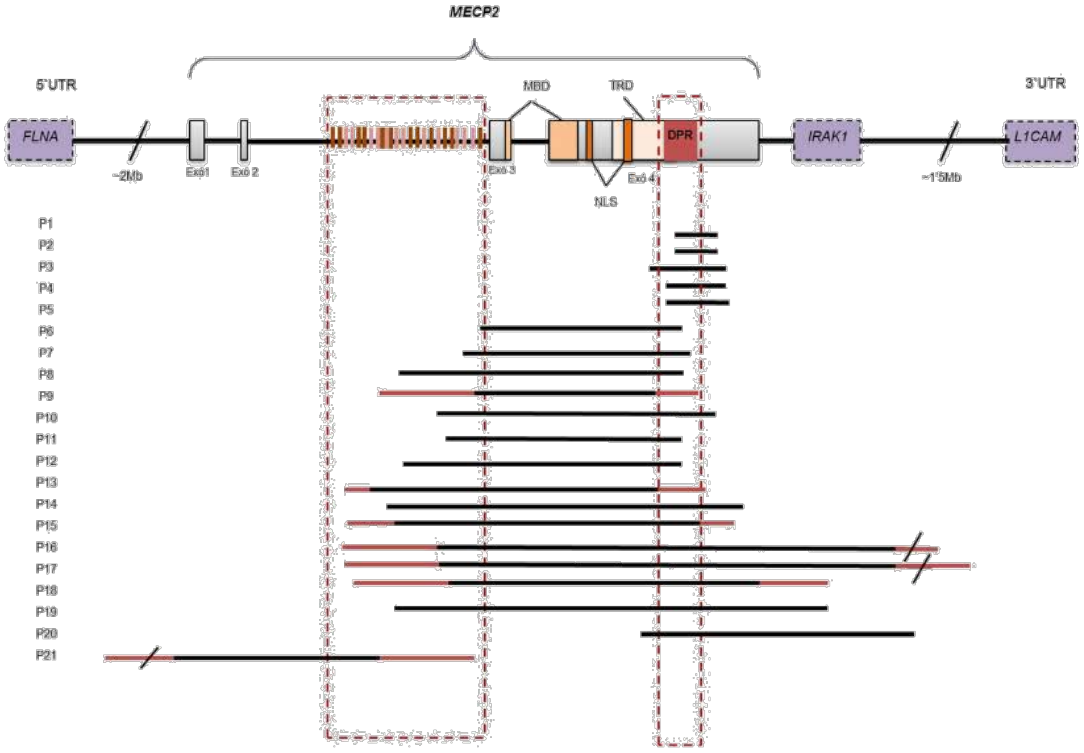


FIGURE 1 Position of the deletions in *MECP2*. Note that there are two regions prone to harbor a breakpoint. The black line indicates the region known with certainty to be deleted; the red line designates the region where only qPCR information is available

3 | RESULTS

For a total of 21 patients without a point mutation or small *indel* detected in the *MECP2* coding region, MLPA was carried out, and at least one exonic probe was missing in each patient (see Supplementary Data S2). These patients were classified depending on the affected exons: only exon 4 of *MECP2* was affected in five patients (P1 to P5), exons 3 and 4 in nine patients (P6 to P15), exons 3, 4, and *IRAK1* gene in four patients (P16 to P19), exon 4 and *IRAK1* in one patient (P20), and exons 1 and 2 in one patient (P21) (see Table 1). All deletions have been confirmed to be de novo.

To validate the MLPA technique, quantitative-PCR (qPCR) was performed in all patients with suspected deletions. qPCR analysis of the respective regions showed results compatible with deletion. Relative ratios of 0.5 ± 0.2 were suggestive of a deletion, whereas ratios of 1 ± 0.2 were indicative of a normal copy number for that region (for more information about the narrowing down of the deletion in each case, see Supplementary Data S3). Several PCR primer sets were evaluated to identify the ones that flank the deletion junction and could amplify, such that Sanger sequencing could be performed. For patients P1, P2, P3, P4, P5, P6, P7, P8, P10, P11, P12, P14, P19, and P20, different pairs of primers (Supplementary Data S1) successfully amplified the junction fragment that was subsequently sequenced (Figure 1). The deletions we have characterized range from 1,235 bp to 85 kb and involve different exons of *MECP2*, sometimes even ending in nearby genes. In case P19, we found a large inversion alongside the deletion.

With the patients' clinical information, we assigned each patient a severity score based on Pineda score. This scoring system gathers information about clinical features for classic RTT such as the patient's age of onset of the first sign, the presence of microcephaly, the ability to sit alone, ambulation, epilepsy, hand use, onset of stereotypies, respiratory function, and language. Complete information of approximately 12 patients was available; three more patients' reports lack one of those aspects, and no information was available about the other six girls. Even if each patient has a unique phenotype, there are some characteristics that are present more frequently among them: 60% of the girls present the first signs before the age of 12 months, 92.3% of the girls have microcephaly, 64.3% have respiratory problems, 93.3% suffers from epilepsy, 93.3% lost their hand use, 80% began with the stereotypies before 36 months of age and 60% before 24 month, and 86.6% lost their language (see Supplementary Data S4). No clear correlation was identified between the size of the deletion in base pairs and the severity of the phenotype, although there seems to be a trend when taking into account the deleted exons: patients with deletions comprising only one exon of *MECP2* have milder symptoms than those with deletions that involve both exons 3 and 4 or contain *IRAK1* as well.

All 21 patients were heterozygous at the *AR* locus and were thus informative for the assay. Patients P2, P4, P8, P9, P15, P16, P17, and P19 have skewed XCI (defined here as ≥80% activity of one X chromosome). However, 13 of 21 girls have an XCI of >70%; even if we do not consider these results completely skewed, they may indicate a cellular trend to inactivate the mutant allele. This phenomenon could account for why some of our patients do not have such a high score as we may have expected for the size of their deletions (for example P15 or P17). Unfortunately, we lack the clinical information about some of the girls with skewed X. The results for each subject are listed in Table 1 along with a summary of all other results obtained.

4 | DISCUSSION

In this study, we have screened a cohort of 21 classic RTT patients with large *MECP2* deletions detected by MLPA. Subsequent qPCR analysis has confirmed the presence of large deletions in all of them. The deletion breakpoints were further characterized by qPCR and long-range PCR with the aim of defining the precise endpoints at the nucleotide level. That last step was achieved in 14 out of 21 patients. The large number of GCs and all the repetitive sequences found in the intronic region of the gene and in the intergenic zones may have increased the difficulty for the polymerase to amplify our targeted products in those cases. Additionally, after characterizing the case of P19 in which an inversion occurred between two different deletions, we cannot dismiss the possibility that more complex rearrangements are present in the genome of those patients interfering with the correct hybridization of the primers. With the introduction of next generation sequencing and specifically, the whole genome sequencing (WGS), the delineation of those unresolved cases could be achieved; although WGS is still not affordable in order to use it as a routine technique for Rett Syndrome patient testing.

Our results showed a wide range of genotypes, from deletions affecting only a single exon to others involving almost the entire *MECP2* gene and the gene located downstream, *IRAK1*. We found only one patient with a deletion in exons 1 and 2 and part of the promoter region of the *MECP2* gene. This is in accordance with previous findings, although a small number of deletions have been reported affecting exons 1 and 2(Archer et al., 2006; Erlandson et al., 2003; Hardwick et al., 2007; Ravn et al., 2005).

Nine patients (P1, P2, P4, P5, P6, P7, P8, P11, and P12) whom we successfully characterized had a breakpoint in the "deletion-prone region" (DPR, GRCh38/hg38 chrX:154,030,619-154,030,770), as defined by Laccone et

al. (2004)(Laccone et al., 2004). Another two patients (P3 and P10) had their breakpoint close to this region (less than 80 bp away). There are two patients (P9 and P13) who could have one of their breakpoints in the DPR as well, but since they are not fully characterized, we cannot confirm this conclusion (see Figure 1). Our finding together with previous studies (Hardwick et al., 2007; Laccone et al., 2004; Ravn et al., 2005; Schollen, Smeets, Deflem, Fryns, & Matthijs, 2003) could better define the junction sequence of the large MECP2 deletions, since 22 of 42 (52.3%) rearrenged alleles have the breakpoint in the DPR. This region is also the hotspot for the smaller deletions (<500 bp) confined within exon 4. The repetitive nature of the DPR has been considered the major cause of genomic instability there; these include the presence of direct and inverted small repeats, the abundance of polypurine residues in the antisense strand and the presence of the χ-sequence GCTGGTGG, which has been found to be highly recombinogenic in the *Escherichia coli* genome (Stahl, Kobayashi, Stahl, & Huntington, 1983). It has been suggested that this sequence stimulates the recombinase BC-dependent system and is responsible of certain deletions that cause human diseases (Amor, Parkert, Globerman, New, & White, 1988; Marshall, Isidro, & Boavida, 1996).

In addition to the DPR, eight patients whom we successfully characterized and seven in whom long PCR failed had a breakpoint in the same intron 2 region (GRCh38/hg38 version chrX:154033244-154052415). The RepeatMasker program (http://www.repeatmasker.org) revealed that 48.9% of this intronic region consists on interspersed repeats, and 17.9% of them are Alu elements. It has been previously hypothesized that those abundant Alu elements interact with the χ-sequence near the DPR making these types of large rearrangements in *MECP2* possible and recurrent (Laccone et al., 2004; Rüdiger, Gregersen, & Kielland-brandt, 1995). Additionally, Alu has proven to be involved in other genomic rearrangements in different genes (Gu et al., 2015; López et al., 2015; Peixoto et al., 2013). The data we provide contribute to strengthen the theory that all these rearrangements do not occur randomly across the gene and its surroundings but in focal areas. Once the deletion is precisely delimited, studies to correct this mutation by CRISPR-Cas9 technology could be considered to regain a complete and functional *MECP2*, among other strategies the cell possess such as homology repair.

We have attempted to establish genotype–phenotype correlations with our patient cohort. Although no clear correlation between the deleted exons and the clinical severity has emerged from this study, we can appreciate some trends (see Figure 2, left). The patient with the deletion involving exons 1 and 2 has a severe phenotype, which seems reasonable because those exons contain the starting sites for both isoforms of *MECP2* and that without that signal, no product could be generated a priori. Patients with a deletion in exon 4 show the mildest phenotype compared to the remaining combinations. This finding can be explained because none of our five deletions in exon 4 occurs in any of the main functional domains of the protein. An exception could be P5 who has the highest score of the group but, in this particular case, the patient was only 2 years old when the score was set, so it can still improve in the following years. Some authors have claimed that deletions involving *IRAK1* generate a more severe phenotype (Hardwick et al., 2007). *IRAK1* is the interleukin 1 receptor-associated kinase and plays a critical role in initiating an innate immune response against foreign pathogens. In our cohort, little difference can be seen when *IRAK1* is added to the deletion but we must admit that the used checklist and scoring system does not take into account the severity or recurrence of the infections of the patients, features that may allow differentiating the effect of having or not *IRAK1* deleted. If we use to determine the severity of the phenotype the Pineda's clinical score, patients reported by Harwick et al. (2007) and patients from our cohort trends to be similar (see Figure 2 right). However, we are aware that exceptions exist and that, sometimes, patients with the same or very similar deletion present a very different phenotype. For example, P4 and P5 have a similar deletion but their score differs by five points. Other examples are the cases described

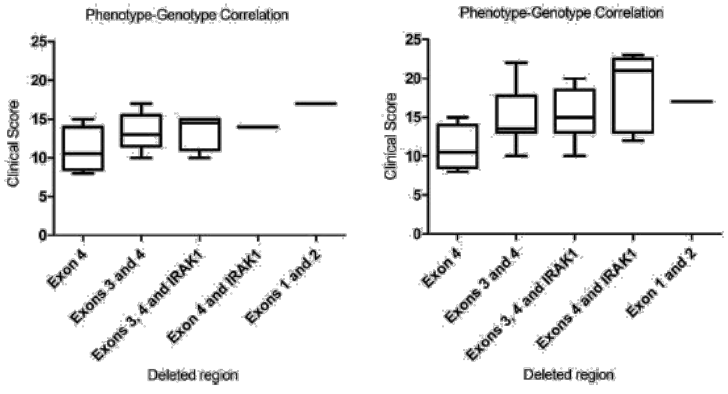


FIGURE 2 Phenotype–genotype correlation according to the deleted region. The left side corresponds to the correlation based on our patients. The right side shows the same correlation based on our patients and the ones reported by Hardwick et al. (2007)

by Bebbington et al. (2012), Mittal, Kabra, Juyal, and BK (2011) or Erlandson et al. (2003). All these cases may suggest another mechanism that alters the direct effect caused by the deletions, possibly a specific methylation pattern causing another molecular alteration in another gene or regulatory domain. Such is the case of the brain-derived neurotrophic factor, *BDNF* (OMIM*113505), gene which is known to protect the carriers of the polymorphism p.Val66Met against early onset epilepsy (Li & Pozzo-Miller, 2014).

Considering the molecular and clinical effect that the lack of a noteworthy part of the coding region of the gene can cause, we were expecting very severe phenotypes. However, the scores in our cohort were not always correlated. The XCI issue has frequently been considered in research on RTT as a potential explanation for the diverse phenotypes generated from the same genotype (Shahbazian, Sun, & Zoghbi, 2002). It has been shown that different cell types can have a different XCI pattern and that the one observed in blood lymphocytes may not be the same as in the brain, the organ in which the majority of symptoms of RTT occur (De Hoon, Monkhorst, Riegman, Laven, & Gribnau, 2015). This phenomenon has accounted for how some females carry a mutation in *MECP2* and are asymptomatic, because of the extreme inactivation of the X chromosome that harbors the aberrant allele (Shahbazian et al., 2002). A similar scenario occurs in mothers carrying a duplication of *MECP2*, who have a skewed XCI and are phenotypically normal unlike their affected offspring who develop *MECP2* duplication syndrome (Van Esch, 2011; Lim, Downs, Wong, Ellaway, & Leonard, 2017). Eight of our patients have a skewed XCI pattern, nine if we lower the threshold to <75% like other authors have done (Hardwick et al., 2007). However, if we lower it to <70% of XCI, four more girls can be included, making a total of 13 patients without a complete random pattern, which could suggest positive selection of cells with an inactivated copy of the defunct *MECP2* allele as a protective mechanism against such large deletions. This hypothesis could explain the relatively mild phenotypes of our cohort. Additionally, we could perform allele specific XCI in two of our patients, P8 and P20 (Personal Data). P8 presents an allele-specific inactivation of 6:94, so most of the mutated allele was inactive, as we expected. Unfortunately, no clinical data were available for this patient. In the case of P20, this technique showed a random inactivation of the gene, so the score might not be so high because no functional domain is present in the deletion and, therefore, the molecular implications for the loss might not be as critical as if they were.

In conclusion, molecular characterization of large rearrangements in *MECP2* is possible in the majority of the cases using the methodology we have exposed. Analysis of that information supports the theory that the 3' end of exon 4 and intron 2 are prone to suffer breaks that can lead to these deleterious big deletions. In addition, gathering clinical data enabled us to define a new set of features that are present in patients with large deletions, such as microcephaly, epilepsy, loss of hand use, loss of language, or onset of stereotypies before 36 months. These data will be very helpful for genetic counseling. A correlation between the severity of the patient and the position of the deletion shows that it is milder when only one exon is deleted and more severe when exons 3 and 4 and *IRAK1* are also involved. In addition, it seems that there is a cellular trend that inactivates the chromosome with the aberrant allele alleviating the final phenotype.

ACKNOWLEDGMENTS

We thank all patients and their families, who contributed to this study. This work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01159), Catalan Association for Rett Syndrome, Fondobiorett and Mi Princesa Rett.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHORS' CONTRIBUTIONS

J.A. and M.P. conceived and supervised the study. S.V., A.P.A., M.R., E G., N.B., and P.P. performed the experiments and collected the data. J.A., S.V., A.P.A., M.R., E.G., N.B., and P.P. analyzed the results. J.A, M.P., and the Rett Working Group provided the patients' samples and clinical and genetic information. A.P.A., S.V., and J.A. wrote the manuscript. All the authors reviewed the article critically for intellectual content.

RETT WORKING GROUP

Hospital Sant Joan de Déu (Barcelona): Maria del Mar O'Callaghan (mocallaghan@sjdhospitalbarcelona.org); Àngels Garcia-Cazorla (agarcia@sjdhospitalbarcelona.org)

Hospital San Borja Arriaran, Santiago de Chile (Chile), Facultad de Medicina Universidad de Chile: Dra Monica Troncoso (monicatroncososch@gmail.com); Dr Guillermo Fariña (guillermofarina@gmail.com)

Hospital Clínico Universitario Virgen de la Arrixaca (Murcia): Dra María Rosario Domingo (mrosario.domingo@gmail.com); Dr. Salvador Ibañez (salibmi@hotmail.com)

Hospital Infantil Universitario Niño Jesús (Madrid): Dr. García Peñas (jgarciadelarape.1961@gmail.com); Dra López Marín (laural.marin@hotmail.com); Dr. González Gutierrez-Solana (luisggsolana@hotmail.com)

Hospital Materno Infantil- Hospital Infanta Cristina (Badajoz): Dr. Enrique Galán Gomez (egalan@unex.es)

Hospital Universitario Fundación Alcorcón (Madrid): Miguel Ángel Martínez Granero (MAMartinezg@fhalcorcon.es)

Hospital Universitario de Girona Dr. Josep Trueta (Girona): Dra María Obon (mobon.girona.ics@gencat.cat)

Parc Taulí Hospital Universitari (Barcelona): Dra Neus Baena Díez (NBaena@tauli.cat)

ORCID

Judith Armstrong https://orcid.org/0000-0003-0588-9307

REFERENCES

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*, *23*, 185–188. https://doi.org/10.1038/13810

Amor, M., Parkert, K. L., Globerman, H., New, M., &White, P. C. (1988). Mutation in the CYP21B gene (Ile-172-%3eAsn) causes steroid 21-hydroxylase deficiency. *Genetics*, *85*, 1600–1604.

Archer, H. L., Whatley, S. D., Evans, J. C., Ravine, D., Huppke, P., Kerr, A., … Clarke, A. J. (2006). Gross rearrangements of the MECP2 gene are found in both classical and atypical Rett syndrome patients. *Journal of Medical Genetics*, *43*, 451–456. https://doi.org/10.1136/jmg.2005.033464

Bebbington, A., Downs, J., Percy, A., Pineda, M., Zeev, B. B., Bahi-Buisson, N., & Leonard, H. (2012). The phenotype associated with a large deletion on MECP2. *European Journal of Human Genetics*, *20*, 921–927. https://doi.org/10.1038/ejhg.2012.34

De Hoon, B., Monkhorst, K., Riegman, P., Laven, J. S. E., & Gribnau, J. (2015). Buccal swab as a reliable predictor for X inactivation ratio in inaccessible tissues. *Journal of Medical Genetics*, *52*(11), 784–790. https://doi.org/10.1136/jmedgenet-2015-103194

Erlandson, A., Samuelsson, L., Hagberg, B., Kyllerman, M., Vujic, M., & Wahlström, J. (2003). Multiplex ligation-dependent probe amplification (MLPA) detects large deletions in the MECP2 gene of Swedish Rett Syndrome patients. *Genetic Testing*, *7*(4), 329–332.

Gu, S., Yuan, B. O., Campbell, I. M., Beck, C. R., Carvalho, C. M. B., Nagamani, S. C. S., … Lupski, J. R. (2015). Alu-mediated diverse and complex pathogenic copy-number variants within human chromosome. *Human Molecular Genetics*, *24*(14), 4061–4077. https://doi.org/10.1093/hmg/ddv146

Hagberg, B., Aicardi, J., Dias, K., & Ramos, O. (1983). A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: Report of 35 cases. *Annals of Neurology*, *14*(4), 471–479. https://doi.org/10.1002/ana.410140412

Hardwick, S. A., Reuter, K., Williamson, S. L., Vasudevan, V., Donald, J., Slater, K., … Christodoulou, J. (2007). Delineation of large deletions of the MECP2 gene in Rett syndrome patients, including a familial case with a male proband. *European Journal of Human Genetics*, *15*, 1218–1229. https://doi.org/10.1038/sj.ejhg.5201911

Kron, M., Howell, C. J., Adams, I. T., Ransbottom, M., Christian, D., Ogier, M., & Katz, D. M. (2012). Brain activity mapping in Mecp2 mutant mice reveals functional deficits in forebrain circuits, including key nodes in the default mode network, that are reversed with Ketamine treatment. *Journal of Neuroscience*, *32*(40), 13860–13872. https://doi.org/10.1523/JNEUROSCI.2159-12.2012

Laccone, F., Jünemann, I., Whatley, S., Morgan, R., Butler, R., Huppke, P., & Ravine, D. (2004). Large deletions of the MECP2 gene detected by gene dosage analysis in patients with Rett Syndrome. *Human Mutation*, *23*, 234–244. https://doi.org/10.1002/humu.20004

Li, W., & Pozzo-Miller, L. (2014). BDNF deregulation in Rett syndrome. *Neuropharmacology*, *76*, 737–746. https://doi.org/10.1016/j.neuropharm.2013.03.024

Lim, Z., Downs, J., Wong, K., Ellaway, C., & Leonard, H. (2017). Expanding the clinical picture of the MECP2 duplication syndrome. *Clinical Genetics*, *91*, 557–563. https://doi.org/10.1111/cge.12814

Liyanage, V. R. B., & Rastegar, M. (2014). Rett syndrome and MeCP2. *NeuroMolecular Medicine*, *16*(2), 231–264. https://doi.org/10.1007/s12017-014-8295-9

López, E., Casasnovas, C., Giménez, J., Matilla-Dueñas, A., Sánchez, I., & Volpini, V. (2015). Characterization of Alu and recombination-associated motifs mediating a large homozygous SPG7 gene rearrangement causing hereditary spastic paraplegia. *Neurogenetics*, *16*, 97–105. https://doi.org/10.1007/s10048-014-0429-6

Marshall, B., Isidro, G., & Boavida, M. G. (1996). Insertion of a short Alu sequence into the hMSH2 gene following a double cross over next to sequences with chi homology. *Gene*, *174*, 175–179. https://doi.org/10.1016/0378-1119(96)00515-X

Mittal, K., Kabra, M., Juyal, R., & BK, T. (2011). De novo deletion in MECP2 in a monozygotic twin pair: A case report. *BMC Medical Genetics*, *12*, 113. https://doi.org/10.1186/1471-2350-12-113

Mnatzakanian, G. N., Lohi, H., Munteanu, I., Alfred, S. E., Yamada, T., MacLeod, P. J. M., … Minassian, B. A. (2004). A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nature Genetics*, *36*(4), 339–341. https://doi.org/10.1038/ng1327

Monrós, E., Armstrong, J., Aibar, E., Poo, P., Canós, I., & Pineda, M. (2001). Rett syndrome in Spain: mutation analysis and clinical correlations. *Brain and Development*, *23*, S251–S253. https://doi.org/10.1016/S0387-7604(01)00374-6

Nan, X., Meehan, R. R., & Bird, A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Research*, *21*(21), 4886–4892. https://doi.org/10.1093/nar/21.21.4886

Neul, J. L., Fang, P., Barrish, J., Lane, J., Caeg, E. B., Smith, E. O., … Glaze, D. G. (2008). Specific mutations in Methyl-CpG-Binding Protein 2 confer different severity in Rett syndrome. *Neurology*, *70*(16), 1313–1321. https://doi.org/10.1212/01.wnl.0000291011.54508.aa

Neul, J. L., Kaufmann, W. E., Glaze, D. G., Clarke, A. J., Leonard, H., Bailey, M. E. S., … Renieri, A. (2010). Rett syndrome: Revised diagnostic criteria and nomenlcature. *Annual Neurology*, *68*(6), 944–950. https://doi.org/10.1002/ana.22124.Rett

Peixoto, A., Pinheiro, M., Massena, L., Santos, C., Pinto, P., Rocha, P., … Teixeira, M. R. (2013). Genomic characterization of two large Alu-mediated rearrangements of the BRCA1 gene. *Journal of Human Genetics*, *58*, 78–83. https://doi.org/10.1038/jhg.2012.137

Ravn, K., Nielsen, J. B., Skjeldal, O. H., Kerr, A., Hulten, M., & Schwartz, M. (2005). Large genomic rearrangements in MECP2.

Human Mutation, *25*(3), 324–329. https://doi.org/10.1002/humu.9320

Rüdiger, N. S., Gregersen, N., & Kielland-brandt, M. C. (1995). One short well conserved region of Alu-sequences is involved in human gene rearrangements and has homology with prokaryotic chi. *Nucleic Acids Research*, *23*(2), 256–260. https://doi.org/10.1093/nar/23.2.256

Scala, E., Longo, I., Ottimo, F., Speciale, C., Sampieri, K., Katzaki, E., … Ariani, F. (2007). MECP2 deletions and genotype-phenotype correlation in Rett syndrome. *American Journal of Medical Genetics, Part A*, *43A*, 2775–2784. https://doi.org/10.1002/ajmg.a.32002

Schollen, E., Smeets, E., Deflem, E., Fryns, J. P., & Matthijs, G. (2003). Gross rearrangements in the MECP2 gene in three patients with Rett syndrome: Implications for routine diagnosis of Rett syndrome. *Human Mutation*, *22*, 116–120. https://doi.org/10.1002/humu.10242

Shahbazian, M. D., Sun, Y., & Zoghbi, H. Y. (2002). Balanced X chromosome inactivation patterns in the Rett syndrome brain. *American Journal of Medical Genetics*, *111*(2), 164–168. https://doi.org/10.1002/ajmg.10557

Singh, J., Saxena, A., Christodoulou, J., & Ravine, D. (2008). MECP2 genomic structure and function: Insights from ENCODE. *Nucleic Acids Research*, *36*(19), 6035–6047. https://doi.org/10.1093/nar/gkn591

Stahl, M. M., Kobayashi, I., Stahl, F. W., & Huntington, S. K. (1983). Activation of Chi, a recombinator, by the action of an endonuclease at a distant site. *Genetics*, *80*, 2310–2313.

Van Esch, H. (2011). MECP2 duplication syndrome. *Molecular Syndromology*, *2*, 128–136. https://doi.org/10.1159/000329580

Vidal, S., Brandi, N., Pacheco, P., Gerotina, E., Blasco, L., Trotta, J.-R., … Armstrong, J. (2017). The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome. *Scientific Reports*, https://doi.org/10.1038/s41598-017-11620-3

Yaron, Y., Ben Zeev, B., Shomrat, R., Bercovich, D., Naiman, T., & Orr-Urtreger, A. (2002). MECP2 mutations in Israel: Implications for molecular analysis, genetic counseling, and prenatal diagnosis in Rett syndrome. *Human Mutation*, *20*(4), 323–324. https://doi.org/10.1002/humu.9069

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Vidal S, Pascual-Alonso A, Rabaza-Gairí M, et al; Rett Working Group. Characterization of large deletions of the *MECP2* gene in Rett syndrome patients by gene dosage analysis. *Mol Genet Genomic Med*. 2019;7:e793. https://doi.org/10.1002/mgg3.793

Capítulo 3

Capítulo 3: Análisis de correlaciones genotipo-fenotipo en el RTT

Más de cincuenta años después de la primera publicación sobre el RTT, y poco menos de dos décadas después del primer informe que lo vincula con el gen *MECP2*, el esfuerzo de la comunidad científica se centra en obtener una mejor comprensión de la compleja genética y biología de esta enfermedad. Los estudios genéticos realizados a partir de la NGS nos han permitido estudiar una mayor cantidad de genes asociados con los fenotipos RTT-*like* simultáneamente, proporcionando un diagnóstico genético para un grupo más amplio de pacientes. Los resultados obtenidos nos han permitido realizar un análisis de correlación genotipo-fenotipo con una gran cohorte de pacientes con RTT y RTT-*like* e intentar esbozar qué vías en la cascada de *MECP2* pueden estar alteradas.

Publicación 4

Título: *The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome.*

Autores: Silvia Vidal, Núria Brandi, Paola Pacheco, Joan Maynou, Guerau Fernandez, Clara Xiol, Ainhoa Pascual-Alonso, Mercè Pineda, *Rett Working Group* and Judith Armstrong.

Referencia: *European Journal of Paediatric Neurology.* 2019 Jul;23(4):609-620.

Resumen:

En este estudio identificamos otros trastornos monogénicos que comparten características solapantes con el RTT. Un total de 40 pacientes con características y diagnóstico clínico de RTT presentaban variantes patogénicas en seis genes asociados con otros trastornos monogénicos. Doce pacientes tenían variantes en *STXBP1*, nueve en *TCF4*, seis en *SCN2A*, cinco en *KCNQ2*, cuatro en *MEF2C* y cuatro en *SYNGAP1*. Estos nuevos hallazgos proporcionan al clínico más información y pistas que podrían ayudar en la prevención de síntomas futuros o en el uso de terapias farmacológicas más adecuadas para cada caso en particular.

EUROPEAN JOURNAL OF PAEDIATRIC NEUROLOGY 23 (2019) 609–620

Official Journal of the European Paediatric Neurology Society

Original article

The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome

S. Vidal [a,d], N. Brandi [b], P. Pacheco [c], J. Maynou [c,d], G. Fernandez [c,d], C. Xiol [a,d], A. Pascual-Alonso [a,d], M. Pineda [a], Rett Working Group, J. Armstrong [c,d,e,*]

[a] Sant Joan de Déu Research Foundation, Barcelona, Spain
[b] School of Medicine, Universitat de Barcelona, Barcelona, Spain
[c] Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain
[d] Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain
[e] CIBER-ER (Biomedical Network Research Center for Rare Diseases), Institute of Health Carlos III (ISCIII), Madrid, Spain

ARTICLE INFO

Article history:
Received 19 December 2018
Received in revised form
12 February 2019
Accepted 28 April 2019

Keywords:
Rett syndrome
Genotype-phenotype correlations
Monogenic disorders
MECP2
RTT
Rett-like

ABSTRACT

Rett syndrome (RTT) is an early-onset neurodevelopmental disorder that is caused by mutations in the MECP2 gene; however, defects in other genes (CDKL5 and FOXG1) can lead to presentations that resemble classic RTT, although they are not completely identical. Here, we attempted to identify other monogenic disorders that share features of RTT. A total of 437 patients with a clinical diagnosis of RTT-like were studied; in 242 patients, a custom panel with 17 genes related to an RTT-like phenotype was run via a HaloPlex-Target-Enrichment-System. In the remaining 195 patients, a commercial TruSight-One-Sequencing-Panel was analysed. A total of 40 patients with clinical features of RTT had variants which affect gene function in six genes associated with other monogenic disorders. Twelve patients had variants in STXBP1, nine in TCF4, six in SCN2A, five in KCNQ2, four in MEF2C and four in SYNGAP1. Genetic studies using next generation sequencing (NGS) allowed us to study a larger number of genes associated with RTT-like simultaneously, providing a genetic diagnosis for a wider group of patients. These new findings provide the clinician with more information and clues that could help in the prevention of future symptoms or in pharmacologic therapy.

© 2019 European Paediatric Neurology Society. Published by Elsevier Ltd. All rights reserved.

* *Corresponding author.* Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Pg. Sant Joan de Déu 2, planta 0, Esplugues de Llobregat, 08950, Barcelona, Spain. Fax: + 34 93 600 9760.
E-mail address: jarmstrong@sjdhospitalbarcelona.org (J. Armstrong).
https://doi.org/10.1016/j.ejpn.2019.04.006
1090-3798/© 2019 European Paediatric Neurology Society. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Rett syndrome (RTT; OMIM#312750) is an early onset neurodevelopmental disorder that almost exclusively affects girls and has an incidence of 1:10,000 live female births.[1] This syndrome is first recognized in infancy with a period of apparently normal development (up to the age of 6–18 months), followed by a regression characterized by a loss or deterioration of speech and purposeful hand use and motor apraxia, which may be associated with epilepsy and dysautonomic features, including disturbed breathing and sleep as well as gastrointestinal dysmotility.[2,3] Hand stereotypies and breathing abnormalities, including hyperventilation and/or breath holding episodes, are distinct clinical features that often present and can help in the diagnosis. Facial dysmorphism is not distinct, and few have subtle dysmorphic facial features that do not enable a clinical diagnosis.[4] The type and severity of symptoms are highly different in individuals.

Although the majority of RTT patients have pathogenic variants in the gene encoding methyl-CpG binding protein 2 (*MECP2*, OMIM*300005), approximately 5% of classic RTT and 25% of atypical RTT patients are negative for *MECP2* pathogenic variants.[5,6] In this group of atypical RTT patients, some have variants in other genes that are also related to RTT, such as cyclin-dependent kinase-like 5 (*CDKL5*; OMIM*300203) and forkhead box protein G1 (*FOXG1*; OMIM*164874). Recently, as a consequence of large-scale genetic screening technologies, other genes not previously related to RTT have been associated with RTT-like phenotypes, such as transcription factor 4 (*TCF4*; OMIM*602272) and myocyte-specific enhancer factor 2C (*MEF2C*; OMIM*60066).[7]

Next generation sequencing (NGS) has emerged as a potentially powerful tool for the study of this type of genetic disease.[8–11] The aim of this study was to continue to extend and improve the diagnosis of monogenic disorders that share features with RTT. Here, we report accurately the 40 cases of a cohort of 437 patients with features of RTT, described by Vidal S. et al. 2017,[9] no *MECP2* defects and variants which affect gene function in six genes associated with other monogenic disorders. Here, we compare different disorders and causative genes to RTT features.

2. Material and methods

2.1. Subjects and ethical issues

A cohort of 437 Spanish patients who presented with clinical features associated with RTT or RTT-like phenotypes was recruited at Sant Joan de Déu Hospital in Barcelona from different Spanish Hospitals.[9] The study was approved by the ethical committees of Hospital Sant Joan de Déu, CEIC: *Comité d'Ètica d'Investigació Clínica* - Fundació Sant Joan de Déu (internal code: PIG-101-15). Patients or their parents gave signed informed consent for genetic studies, and blood samples from patients and controls were obtained according to the Helsinki Declaration of 1964, as revised in 2004.[12] Patients had been diagnosed following the usual clinical parameters[13] and according to the recently revised RTT Search International Consortium criteria and nomenclature.[6] Patients who almost completely fulfilled the criteria, including the main features, such as psychomotor delay with or without regression stereotypic hand movements and absent language or limited to only a few words, were also included.

2.2. Library preparation and bioinformatic pipeline

Libraries of the 242 patients' samples were generated using a custom-made panel with 17 genes associated with a RTT-like phenotype through HaloPlex Target Enrichment System (Agilent Technologies, Santa Clara, CA) and 195 using the TruSight One Sequencing Panel kit (Illumina, San Diego, CA) (Supplementary Table S1)[9]; both according to the manufacturer's sample preparation protocol, and all of them were sequenced on an Illumina NextSeq 500. The variant calling pipeline was developed at the Bioinformatics Unit from the Molecular Genetics Department at the Sant Joan de Déu Hospital. The bioinformatic analysis was divided into several steps: quality control, alignment, variant calling, variant annotation and, finally, filtering. Before and after the adaptor and low quality reads were removed (cutadapt v.1.13), read quality control was assessed using FastQC v.0.11.5.[14,15] The reads were aligned to the human reference genome sequence (hg19/GRCh37) using Burrows-Wheeler Aligned through BWA-MEM v.0.7.15.[16] The aligned reads were filtered by means of mapping quality and duplicates to ensure high quality data using BEDtools v.2.26.0[17] and Picard tools v2.9.0.[18] Once the sequences were filtered, variant calling was determined using SAMTools v.1.5,[19] FreeBayes v1.1.0,[20] VarScan v2.4.0[21] and GATK v3.7.[22] The variants were annotated using SnpEff v.4.3 and included nucleotide and amino acid annotations (dbSNP), population frequencies (gnomAD and internal database), and clinical information (Clinvar, OOMIM). The variants were filtered based on population frequencies (<0.01), coverage (>= 20), amino acid impact (High or Moderate according to SnpEff), pathogenic scores, clinical significance and inheritance patterns. To reduce the amount of variants to be analysed, another filtering layer that took into account specific lists of genes that contained putative targets (RTT-like genes, GABA and glutamate pathway and epilepsy genes) was applied (Supplementary Table S1). Copy-number variant (CNV) were detected using R-package ExomeDepth v1.1.10 based on read-depth method. Deletions and duplications identified were annotated using Database of Genomic Variants version March 2016 and internal database,[23] all CNV detected by NGS of them were confirmed by CGH array 400k.

2.3. Molecular analysis

To identify the potential causative variants, we checked the variants by matching their affected phenotypes and inheritance patterns of respective genes checked by Sanger sequencing of the index cases and their parents. Moreover, we considered the pathogenicity by *in silico* predictors:

MutationTaster (http://www.mutationtaster.org/), SIFT-PROVEAN (http://sift.jcvi.org/), Sorts Intolerant From Tolerant (SIFT), Protein Variation Effect Analyzer (PROVEAN) and Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/), and American College of Medical Genetics and Genomics (ACMG) guidelines were used (https://www.acmg.net/docs/standards_guidelines_for_the_interpretation_of_sequence_variants.pdf). In addition, databases in the literature, such as the RettBASE: *MECP2* Variation Database, the Exome Aggregation Consortium (ExAC), Genome Aggregation Database (gnomAD), HGMD® Professional 2018.2 and the Single Nucleotide Polymorphism Database (dbSNP), were revised. All variants detected and their subsequent segregation studies were performed by Sanger sequencing.

3. Results and discussion

A total of 40 patients with clinical features of RTT-like had variants which affect gene function in six genes associated with other monogenic disorders (Fig. 1 and Table 1). All variants were not registered in ExAC or gnomAD. A total of 25% of these variants (10/40) were considered pathogenic or likely pathogenic in ClinVar, and only in 10% of variants (4/40) did one or more programs predict that there could be possibly benign or neutral changes. The variants information is shown in Supplementary Table S2.

We obtained partial or complete clinical information from 37 patients (Table 2); therefore, the data and percentages are related to the total, with the clinical information of the specific feature available. All the patients presented with psychomotor delay and mental retardation, and these were severe in 75.7% of the patients (28/37). The most common characteristics of our cohort of RTT-like patients were autistic features (30/34) and breathing dysfunction (30/34). A total of 51.7% of the patients (15/29) showed acquired microcephaly, and 34.4% (11/32) had abnormalities of the brain revealed by magnetic resonance imaging (MRI). A total of 74.2% of patients (23/31) presented with hypotonia, and 68.4% (26/38) were able to walk, while a further ten patients (26.3%) who had a more severe phenotype never walked, and two of them (5.3%) had lost this ability. Language skills were limited to a few words in 25.8% of patients (8/31), one (3.2%) had lost the skills, and 75.7% (22/31) never acquired them. Additionally, epilepsy was present in 61.1% of the patients (22/36), 79.4% (27/34) showed typical RTT hand stereotypies, hand wringing or hand washing, and 53.3% of patients (16/30) showed mood disturbances the most common of which was disruptive agitation (14/30); only 13.3% showed sudden laughing (4/30) and 6.7% aggressiveness (2/30).

In Table 3, we summarize the most relevant clinical manifestations of the patients with typical RTT associated with the classic *MECP2*, *CDKL5* and *FOXG1* alterations and those observed in the present series of Rett-like associated genes. Many overlapping characteristics were detected; some cardinal features are also highlighted, such as the absence of developmental regression in patients with pathogenic variants at *TCF4*, *KCNQ2* and *MEF2C* or the absence of breathing dysfunction in those with molecular alterations at *STXBP1*, *SCN2A*, *KCNQ2*, *MEF2C* or *SYNGAP1*.

3.1. *STXBP1 gene*

We describe twelve RTT-like patients with different ten variants in the *STXBP1* gene (Syntaxin-binding protein 1; OMIM#602926). This protein plays an important role in presynaptic vesicle docking and fusion, a necessary mechanism for neurotransmitter release.[24] Reduced *STXBP1* expression has been shown to increase synaptic depression at both GABAergic and glutamatergic synapses, with greater impact on GABAergic interneurons,[25] a pathway that is also altered in RTT patients.[26] Mutations in *STXBP1* have been associated with EEIE7 and a series of neurodevelopmental disorders, including RTT-like syndrome. Our patients showed a combination of stereotypies, autistic features, and regression that have been already described in patients with mutations in *STXBP1*.[27]

The twelve *STXBP1* variants identified in our RTT-like patients include one in-frame deletion, one duplication, two splicing site variants and six missense changes, all within the protein domains. Eight variants were demonstrated to be *de novo*, and one was inherited from the mother; in one case, the mother was not a carrier, and no sample was available from the father; for two cases, inheritance information was not available.

Regarding missense variants, the previously described *STXBP1* variants p.R292C (x2), p.R292H (x2), p.R406C and c.1359+1G > T were *de novo* and have been related to EEIE7.[28] Within the three novel variants, p.R235Q was found to be *de novo*, p.S516R was inherited from an asymptomatic mother and, for p.K7E, the family study was not available. All pathogenicity prediction tools predicted that these variants were likely to be pathogenic. Both splicing variants can affect the correct splicing of the RNA, c.326-3C > G mutates the splice acceptor site, and c.1359+1G > T disrupts the donor splice site. Finally, if the duplication of exon 10 is in tandem, it could affect the structure of the protein and its functionality. To confirm the pathogenicity of these variants, functional studies should be done. For the inherited variants, studies should be performed to see if the carrying progenitor is a mosaic and whether the RNA of these patients is altered.

We had partial or complete clinical information about all patients, except patient 11 (c.1359+1G > T). We found that severe intellectual disability (ID) with epilepsy and absence of speech are the most common clinical features present in patients with likely pathogenic variants in *STXBP1*. Most of them had severe to profound ID, and only one, patient 7 (seizure-free), had a moderate ID. Nearly 95% of patients published to date present with epilepsy. In our cohort, 8 out of 11 (73%) presented with epilepsy.[29] There were only three patients without seizures: patient 3 c.326-3C > G, patient 9 with duplication and patient 7 with p.R292H. As in patient 8 (also carrying the p.R292H), the onset of epilepsy occurred at 7 years, and patient 7 was only 3 years old; we therefore hypothesized that epilepsy may still appear and that this feature may be related to the affected variant. In fact, when comparing the four patients with changes in arginine 292, it

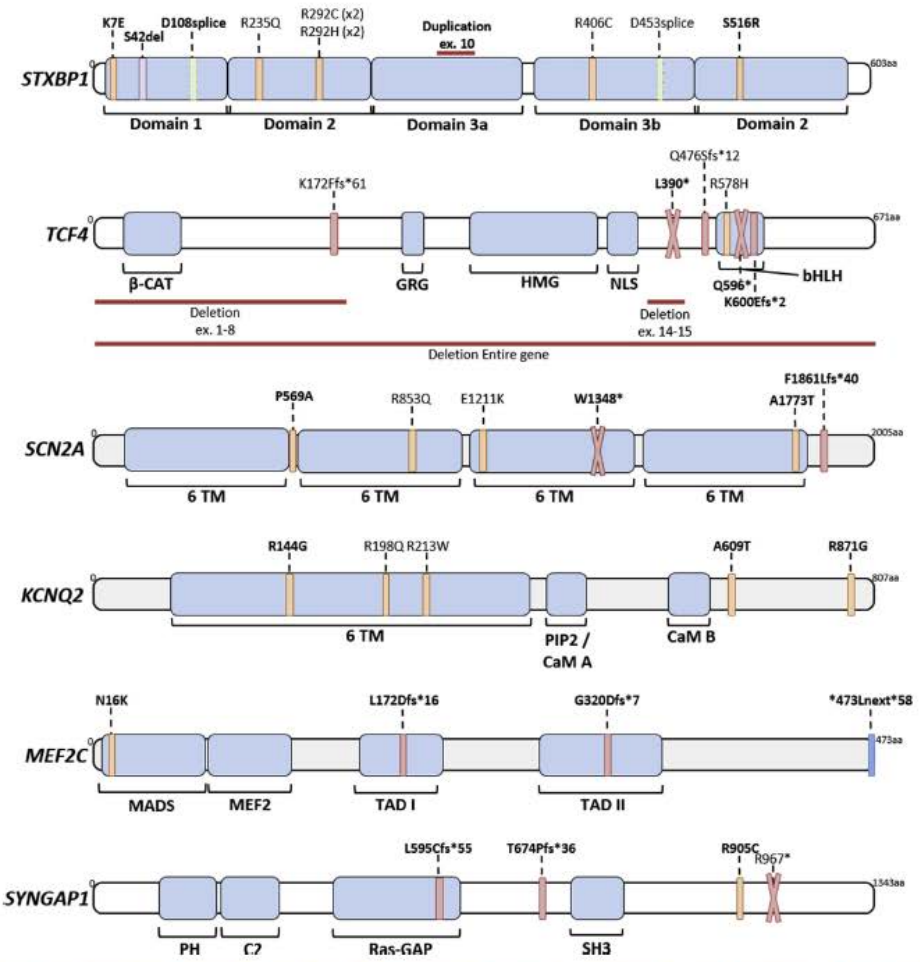


Fig. 1 – Protein structure and location of the variants identified in RTT-like patients. Novel variants in bold; orange box, missense variants; red box, frameshift deletions; red cross, nonsense variants; green box, splice site variants; blue box, stop lost variant; and purple box, in-frame deletion. Abbreviations: β -CAT = N-terminal β-catenin-binding domain; GRG = the interaction domain with the groucho/TLE transcriptional co-repressors; HMG = DNA-binding domain; NLS = the nuclear localization signal; bHLH = basic helix-loop-helix; 6TM = six transmembrane domains; PIP2 = proximal C-terminal domain that binds phosphatidylinositol 4,5-bisphosphate; CaM A = domain which binds calmodulin; CaM B = the more distal domain which binds calmodulin; MADS-box and MEF2 = N-terminal region involved in DNA binding and dimerization; TADI and TADII = activation domains; PH = Pleckstrin Homology domain; C2 = C2 domain; SH3 = SRC Homology 3 domain.

seems that the change at this residue to a cysteine is slightly more severe than the change to histidine. Although the number of patients was very small, we observed that the two patients with a change to cysteine presented with earlier onset of epilepsy, were microcephaly and had not yet acquired walking/ambulation. In contrast, the other two patients with the histidine alteration were able to walk. Regarding the other patient without epilepsy (patient 9), we must conduct further studies to determine whether the duplication of exon 10 of the gene is causative of disease. First, it is important to study the segregation to know if the variant was inherited from parents or is *de novo*. Then, it must be determined if the duplication is in tandem and whether it negatively affects the reading frame of the protein, creating a non-functional protein.

3.2. TCF4 *gene*

We found nine RTT-like patients with anomalies in the *TCF4* gene: three gross deletions, one in-frame indel, three frameshift, one nonsense and one missense change. Six of these

Table 1 – Potentially pathogenic and causative variants detected. Abbreviations: NA = Not Available; Path = Pathogenic; DC = Disease causing; Dam = Damaging; Del = Deleterious; Pol = Polymorphism; Tol = Tolerate; Neu = Neutral; ACMG = American College of Medical Genetics; PA = Potential Alteration; NE = No effect; PVS = Pathogenic Very Strong; PM = Pathogenic Moderate; PP = Pathogenic Supporting; BP = Benign Supporting.

Gene	Patient	Inheritance	Zygosity	Type of seq. Change	cDNA/Protein change	dbSNP	ClinVar	Mutation taster, PROVEAN, SIFT, Polyphen-2, HSF3.1	ACMG Classification
STXBP1	1	NA	Heter.	Missense	c.19A > G/p.(K7E)	–	–	DC, Dam, Del, Dam, PA	PM2, PP3
(NM_003165.3)	2	*de novo*	Heter.	In-frame deletion	c.128_130delTCC/p.(S42del)	–	–	DC, -, -, -, -	PM2, PM4
	3	Mother not carrier	Heter.	Splicing variant	c.326-3C > G/-	–	–	-, -, -, -, Br Acceptor	PM2, PP3
	4	*de novo*	Heter.	Missense	c.704G > A/p.(R235Q)	–	–	DC, Dam, Del, Dam, PA	PM2, PP3
	5	*de novo*	Heter.	Missense	c.874C > T/p.(R292C)	–	–	DC, Dam, Del, Dam, PA	PM2, PM5, PP3
	6	*de novo*	Heter.	Missense	c.874C > T/p.(R292C)	–	–	DC, Dam, Del, Dam, PA	PM2, PM5, PP3
	7	*de novo*	Heter.	Missense	c.875G > A/p.(R292H)	rs796053361	Likely path. Allele	DC, Dam, Del, Dam, PA	PM2, PM5, PP3, PP5
	8	*de novo*	Heter.	Missense	c.875G > A/p.(R292H)	rs796053361	Likely path. Allele	DC, Dam, Del, Dam, PA	PM2, PM5, PP3, PP5
	9	NA	Heter.	Gross duplication	129pb incl. ex. 10/-	–	–	-, -, -, -, -	–
	10	*de novo*	Heter.	Missense	c.1216C > T/p.(R406C)	rs796053367	Likely path. Allele	DC, Dam, Del, Dam, PA	PM2, PM5, PP3, PP5
	11	*de novo*	Heter.	Splicing variant	c.1359+1G > T/-	–	–	-, -, -, -, Br Donor	PVS1, PM2, PP3
	12	Maternal	Heter.	Missense	c.1548C > A/p.(S516R)	–	–	DC, Dam, Del, Dam, PA	PM2, PP3
TCF4	13	*de novo*	Heter.	Frameshift deletion	c.514_517delAAAG/p.(K172Ffs*61)	rs398123561	Path. Allele	DC, -, -, -, -	PVS1, PM2, PP5
(NM_001083962.1)	14	*de novo*	Heter.	Gross deletion	c.1069 + 118_1350 + 119del5450/-	–	–	-, -, -, -, -	–
	15	*de novo*	Heter.	In-frame indel	c.1169_1175delinsAAA/p.(L390*)	–	–	DC, -, -, -, -	PVS1, PM2
	16	*de novo*	Heter.	Frameshift deletion	c.1438delC/p.(Q476Sfs*12)			DC, -, -, -, -	PVS1, PP2, PP3
	17	*de novo*	Heter.	Missense	c.1733G > A/p.(R578H)	rs121909123	Path. Allele	DC, Dam, Del, Dam, NE	PM1, PM2, PM5, PP3, PP5
	18	*de novo*	Heter.	Nonsense	c.1786C > T/p.(Q596*)	–		DC, -, -, -, -	PVS1, PM1, PM2, PP3
	19	NA	Heter.	Frameshift deletion	c.1798_1799delAA/p.(K600Efs*2)	–	–	DC, -, -, -, -	PVS1, PM1, PM2
	20	NA	Heter.	Gross deletion	2.2mb incl. entire gene/-	–	–	-, -, -, -, -	–
	21	NA	Heter.	Gross deletion	0.42mb incl. ex.1–8/-	–	–	-, -, -, -, -	–
SCN2A	22	Paternal	Heter.	Missense	c.1705C > G/p.(P569A)	–	–	DC, Tol, Del, Dam, PA	PM2, PP2, PP3
(NM_021007.2)	23	*de novo*	Heter.	Missense	c.2558G > A/p.(R853Q)	rs794727152	Path. Allele	DC, Dam, Del, Dam, PA	PM2, PP2, PP3, PP5
	24	Mother not carrier	Heter.	Missense	c.3631G > A/p.(E1211K)	rs387906684	Path. Allele	DC, Dam, Del, Dam, PA	PM2, PP2, PP3, PP5
	25	*de novo*	Heter.	Nonsense	c.4043G > A/p.(W1348*)	–	–	DC, -, -, -, -	PVS1, PM1, PM2, PP3
	26	*de novo*	Heter.	Missense	c.5317G > A/p.(A1773T)	rs796053162	Likely path. Allele	DC, Dam, Del, Dam, NE	PM2, PM5, PP2, PP3, PP5
	27	*de novo*	Heter.	Frameshift deletion	c.5583delT/p.(F1861Lfs*40)	–	–	DC, -, -, -, -	PVS1, PM2, BP4
KCNQ2	28	NA	Heter.	Missense	c.430C > G/p.(R144G)	–	–	DC, Dam, Del, Dam, PA	PM2, PM5, PP2, PP3, PP5
(NM_172107.3)	29	*de novo*	Heter.	Missense	c.593G > A/p.(R198Q)	rs796052621	Conflicting path.	DC, Dam, Del, Dam, PA	PM1, PM2, PP2, PP3
	30	NA	Heter.	Missense	c.637C > T/p.(R213W)	rs118192203	Path. Allele	DC, Dam, Del, Dam, PA	PM1, PM2, PM5, PP2, PP3
	31	Maternal/ paternal	Hom.	Missense	c.1825G > A/p.(A609T)	–	–	Pol, Tol, Neu, Ben, NE	PM2, PP2
	32	Maternal	Heter.	Missense	c.2611A > G/p.(R871G)	–	–	DC, Dam, Neu, Ben, PA	PM2, PP2, PP3
MEF2C	33	*de novo*	Heter.	Missense	c.48C > G/p.(N16K)	–	–	DC, Dam, Del, Dam, PA	PM1, PM2, PP3
(NM_002397.4)	34	*de novo*	Heter.	Frameshift insertion	c.513_514insGA/p.(L172Dfs*16)	–	–	DC, -, -, -, -	PVS1, PM2
	35	*de novo*	Heter.	Frameshift deletion	c.959_960delGT/p.(G320Dfs*7)	–	–	DC, -, -, -, -	PVS1, PM2
	36	*de novo*	Heter.	Stop loss	c.1421G > T/p.(*473Lnext*58)	–	–	Pol, -, -, -, -	PM2, PM4

(continued on next page)

Table 1 – (continued)

Gene	Patient	Inheritance	Zygosity	Type of seq. Change	cDNA/Protein change	dbSNP	ClinVar	Mutation taster, PROVEAN, SIFT, Polyphen-2, HSF3.1	ACMG Classification
SYNGAP1 (NM_006772.2)	37	Mother not carrier	Heter.	Frameshift deletion	c.1783delC/p.(L595Cfs*55)	rs587780470	Likely path. Allele	DC, -, -, -, -	PVS1, PM2
	38	*de novo*	Heter.	Frameshift deletion	c.2020delA/p.(T674Pfs*36)	–	–	DC, -, -, -, -	PVS1, PM2
	39	NA	Heter.	Missense	c.2713C > T/p.(R905C)	–	–	DC, Dam, Del, Dam, -	PM2
	40	NA	Heter.	Nonsense	c.2899C > T/p.(R967*)	–	–	DC, -, -, -, -	PVS1, PM2

variants are novel: c.1069 + 118_1350 + 119del5450, c.1169_1175delinsAAA, c.1786C > T (p.Q596*), c.1798_1799delAA (p.K600Efs*2), deletion of 2.2 Mb (the entire gene) and deletion of 0.42 Mb (including exons 1–8). *TCF4* encodes a broadly expressed basic helix-loop-helix (bHLH) protein that forms a homodimer or heterodimer with other bHLH proteins. These dimers bind DNA at Ephrussi (E) box sequences. Alternative splicing produces a number of different *TCF4* isoforms with distinct N-termini that differ in their subcellular localization and transactivation capacity.[30] Seven out of the nine variants found produced a truncated protein, leading to haploinsufficiency of the transcription factor. The gross deletion encompassing the whole gene also results in haploinsufficiency of the protein. In addition, we found a missense variant in the bHLH domain located in a recurrent mutation site. This variant affects an evolutionarily highly conserved arginine residue, constituting the E-box recognition motif. It has been previously demonstrated that such an impairment of the functional bHLH domain reduces the interaction with *ASCL1* in transactivating an E-box-containing reporter construct to a similar degree as that of haploinsufficient stop mutations.[31]

Mutations in *TCF4* have been associated with Pitt-Hopkins syndrome (PTHS; OMIM*610954)that is characterized by ID, epilepsy, microcephaly, facial dysmorphisms, postnatal growth restriction, and intermittent hyperventilation.[31] Episodic hyperventilation/apnea, microcephaly, and autism spectrum disorders (ASD)-related stereotypies hand movements may steer clinicians towards a misdiagnosis of RTT-like rather than PTHS. The presence of distinct facial features is more consistent with PTHS and helps to distinguish PTHS from RTT, but patients do not always have these facial distinctions, which are often not clearly defined during the first year of life.[32]

We had partial or completed clinical information of all *TCF4* RTT-like patients. Interestingly, abnormal MRI (4/7), absence of walking (4/9) or loss of walking ability (1/9), absence of speech (7/8), seizures (3/9) and autism features (8/8) were observed in our patients, suggesting a possible RTT-like phenotype. There was no correlation between the clinical features and the variants. Patients 13 and 16, carrying the upstream frameshift (p.Lys172Phefs*61, supposedly the more aggressive variant) and the missense variant (p.R578H, supposed to be the less affected protein), respectively, had the most moderate clinical characteristics: purposely hand use, ability to walk and only moderate mental retardation. Seizures were not present in the patient with the missense variants, which contradicts what had been previously suggested.[33]

3.3. SCN2A *gene*

SCN2A (neuronal voltage-gated sodium channel NaV1.2; OMIM*182390) encodes one of the sodium channels involved in the initiation and propagation of action potentials in numerous neuron classes. *SCN2A* is expressed early in brain development; each domain of the protein contains six membrane-spanning segments S1–S6, where S1–S4 forms the voltage-sensing domains, and S5–S6 forms the pore loops and DEKA-selectivity filter. Variants in *SCN2A* are associated with three disorders: gain-of-function variants leading to

EUROPEAN JOURNAL OF PAEDIATRIC NEUROLOGY 23 (2019) 609–620

Table 2 – Summary of clinical information available for all patients with mutations in the STXBP1, TCF4, SCN2A, KCNQ2, MEF2C and SYNGAP1 genes. Abbreviations: NA = Not available; y = years; m = month; AED = Antiepileptic drugs; VPA = Valproate; CZ = Carbamazepine; LEV = Levetiracetam; ETO = Etosuximide; LTG = Lamotrigine; ACTH = adrenocorticotrophin hormone; CLB = Clobazam; CLN = Clonazepam; ZNS = Zonisamide; PRM = Primidone; OXC = Oxcarbazepine; PHT = Phenytoin; RFM = Rufinamide; BRV = Brivaracetam.

Patient	Gene	Sex/Age evaluation	Acquired microcephaly	Brain MRI	Hypotonia	Psychomotor delay	Walking/Age	Apraxic gait	Speech	Mental retardation	Autistic behaviours	Breathing dysfunction	Hands use
P1	STXBP1	F/11y	No	Normal	No	Yes	Aided/22m	No	Loss	Profound	Yes	Yes	Yes
P2	STXBP1	F/9y	Yes	Normal	Yes	Yes	Aided/NA	Yes	NA	Moderate	No	No	Yes
P3	STXBP1	M/18y	NA	Normal	Yes	Yes	Aided with help/NA	No	No	Profound	Yes	No	NA
P4	STXBP1	M/13y	No	Normal	No	Yes	Aided/2y 3m	No	NA	Profound	Yes	No	Yes
P5	STXBP1	M/6y	Yes	Abnormal	Yes	Yes	No/-	–	No	Profound	Yes	Yes	Yes
P6	STXBP1	M/14y	Yes	Abnormal	Yes	Yes	No/-	–	No	Profound	Yes	No	Yes
P7	STXBP1	F/4	No	Abnormal	No	Yes	Aided with help/3y 9m	Yes	No	Moderate	No	No	Yes
P8	STXBP1	M/8y	No	Normal	Yes	Yes	Aided with help/NA	Yes	No	Profound	NA	No	No
P9	STXBP1	F/32y	Yes	Normal	No	Yes	Aided/2y	No	No	Profound	Yes	No	No
P10	STXBP1	F/5y	No	NA	Yes	Yes	Loss/NA	–	No	Profound	Yes	NA	NA
P11	STXBP1	M/14y	NA	NA	NA	NA	Aided/NA	Yes	NA	NA	NA	NA	NA
P12	STXBP1	F/9y	No	Abnormal	No	Yes	Aided/2y	Yes	No	Profound	Yes	Yes	Yes
P13	TCF4	F/9y	No	Normal	No	Yes	Aided/30m	Yes	No	Moderate	Yes	Yes	Yes
P14	TCF4	F/15y	Yes	Normal	Yes	Yes	Aided with help/NA	Yes	No	Profound	Yes	No	No
P15	TCF4	F/17y	Yes	NA	Yes	Yes	Loss/NA	–	NA	Profound	Yes	NA	NA
P16	TCF4	F/12y	NA	Abnormal	No	Yes	Aided/NA	No	No	Moderate	Yes	Yes	No
P17	TCF4	F/13y	Yes	Abnormal	Yes	Yes	Aided/3y	Yes	Few words	Profound	Yes	Yes	Yes
P18	TCF4	F/6y	Yes	NA	Yes	Yes	No/-	–	No	Profound	Yes	NA	NA
P19	TCF4	F/2y 4m	No	Normal	Yes	Yes	No/-	–	No	Profound	Yes	No	NA
P20	TCF4	F/7y	NA	Abnormal	NA	Yes	No/-	–	No	Profound	Yes	No	Yes
P21	TCF4	M/1y 6m	No	Abnormal	No	Yes	No/-	–	No	Profound	NA	No	Yes
P22	SCN2A	M/5y	Yes	NA	NA	Yes	No/-	–	No	Profound	Yes	No	NA
P23	SCN2A	M/6y	No	Normal	NA	Yes	No/-	–	No	Profound	Yes	No	No
P24	SCN2A	F/NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
P25	SCN2A	F/5y	NA	Normal	NA	Yes	Aided/NA	NA	Few words	Profound	Yes	No	Yes
P26	SCN2A	F/25y	Yes	Normal	Yes	Yes	Aided/3y	Yes	No	Profound	Yes	Yes	Yes
P27	SCN2A	F/25y	Yes	Normal	NA	Yes	Aided/2y	NA	Few words	Profound	Yes	NA	Yes
P28	KCNQ2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
P29	KCNQ2	F/5y	Yes	Abnormal	Yes	Yes	No/-	No	Few words	Profound	Yes	Yes	Yes
P30	KCNQ2	M/5y	No	Normal	Yes	Yes	Aided/6y	No	NA	Moderate	NA	No	No
P31	KCNQ2	F/4y	No	Normal	Yes	Yes	Aided/24m	No	No	Profound	Yes	No	Yes
P32	KCNQ2	M/2y	Yes	Normal	Yes	Yes	No/-	–	No	Profound	Yes	No	Yes
P33	MEF2C	F/24y	NA	Normal	Yes	Yes	Aided with help/6y	NA	Few words	Profound	Yes	NA	NA
P34	MEF2C	F/6y	NA	Abnormal	Yes	Yes	Aided/NA	NA	Few words	Profound	No	NA	NA
P35	MEF2C	F/8y	NA	Normal	Yes	Yes	Aided with help/1y 2m	NA	No	Profound	Yes	NA	NA
P36	MEF2C	F/18y	NA	Abnormal	Yes	Yes	Aided/3y	NA	No	Profound	Yes	No	NA
P37	SYNGAP1	F/7y 8m	No	Normal	Yes	Yes	Aided/22m	Yes	Few words	Moderate	No	No	Yes
P38	SYNGAP1	F/17y	Yes	NA	NA	Yes	Aided/NA	Yes	Few words	Moderate	Yes	NA	NA
P39	SYNGAP1	F/20y	Yes	Normal	Yes	Yes	Aided/4y 6m	Yes	NA	Moderate	Yes	No	Yes
P40	SYNGAP1	F/4y	No	Normal	Yes	Yes	Aided/3y 6m	Yes	NA	Moderate	Yes	No	Yes

Table 3 – Genes and phenotypic overlap with RTT. Abbreviations: EEIE = epileptic encephalopathy infantile-onset; PTHS = Pitt-Hopkins syndrome; MR = mental retardation; XLD = X-linked dominant; AD = autosomal dominant. Grey boxes are clinical characteristics that are not common with RTT.

	MECP2 MIM#312750	CDKL5 MIM#300672	FOXG1 MIM#613454	STXBP1 MIM#612164	TCF4 MIM#610954	SCN2A MIM#613720	KCNQ2 MIM#613720	MEF2C MIM#613443	SYNGAP1 MIM#612621
	RTT	EEIE2	RTT, congenital	EEIE4	PTHS	EEIE11	EEIE7	MR20	MR5
Developmental regression	+	+	+	+	–	+	–	–	+
Speech deficit/lost	+	+	+	+	+	+	+	+	+
Gait abnormalities	+	+	+	+	+	+	+	+	+
Hand use lost/absent	+	+	+	+	+	+	+	+	+
Stereotypic movements	+	+	+	+	+	+	+	+	+
Breathing dysfunction	+	+	+	–	+	–	–	–	–
Learning impairment	+	+	+	+	+	+	+	+	+
Seizures	+	+	+	+	+	+	+	+	+
Microcephaly	+	+	+	+	+	+	–	–	+
Dysmorphic facial features	–	+	–	–	+	–	–	+	–
Inheritance	XLD	XLD	AD	AD	AD	AD	AD	AD	AD

infantile-onset epileptic encephalopathy-11 (EEIE11; OMIM*613721) and benign familial infantile seizures-3 (BFIS3, OMIM*607745) and variants with diminished channel activity that leads to ASD/ID.[34]

We found six RTT-like patients carrying variants in the *SCN2A* gene: four missense, one nonsense and one frameshift variants. Two of them (p.R853Q and p.E1211K) have been associated with West syndrome and neonatal-infantile seizures.[35] Regarding the novel variants, the missense variant p.A1773T is predicted to be pathogenic by *in silico* analyses; in fact, another reported variant that affects the same codon (p.A1773V) has been related to ASD/ID.[36] Pathogenicity of the missense variant p.P569A is difficult to assess; despite the fact that it was considered damaging by 1 of 4 predictors, it had been inherited from the father. Moreover, this variant is located between the two first transmembrane domains in a region with unknown functional implications. Therefore, further studies are needed to define the effects of this variant. Although the frameshift variant is at the C-terminal end, outside the transmembrane regions, there are other frameshift deletions described in this region that are related with EEIE and ASD/ID.[37]

We had clinical information for five patients. We found that 3 of the 5 patients had a very early onset of seizures (neonatal or early infancy), as already reported by Kong et al.[38] The patient with the recurrent p.R853Q variant presented epilepsy onset at 3 months of life, similar to other patients described in the literature.[36] We did not find any of the abnormalities described by Kong and colleagues in brain MRI scans of four of our patients. Other patient series have found that variants located outside of the transmembrane domains were more likely to cause a severe phenotype.[35] However, Kong et al. did not find this correlation; in our limited cohort, we did not find this correlation either.

3.4. *KCNQ2 gene*

We described five RTT-like patients with five different missense variants in the *KCNQ2* gene (potassium voltage-gated channel subfamily Q, member 2; OMIM#602235). In neuronal cells, *KCNQ2* and *KCNQ3* heterotetramerize to give rise to the M current (IM), a key player for the regulation of neuronal excitability.[39] Variants in this gene are responsible for a wide phenotypic spectrum of epileptic diseases, ranging from infantile-onset epileptic encephalopathy-7 (EIEE7, OMIM*613720) to benign familial neonatal seizures-1 (BFNS1, OMIM*121200).[40]

Two of the identified variants (p.A609T and p.R871G) are located outside of the known protein domains. The p.A609T variant is the only homozygous change identified in the present study that had been inherited from carrier parents (consanguineous family), and it is located close to the calmodulin binding distal domain. The 4 *in silico* predictors consider p.A609T as a benign variant since, in two of the alternative transcripts but not in the canonical one, it is a synonymous change. Though p.A609T seems to be benign, in our opinion, the absence of this variant in the control population and its homozygosity suggests that it could be producing a protein malfunction. Functional studies are needed to validate our hypothesis. Furthermore, p.R871G is located near

the C-terminal of the protein, and, as a missense variant, it is predicted to be pathogenic by *in silico* analyses; in fact, another reported variant affecting the same codon (p.R871S) has been related to EEIE7.[41] The other three variants are located in the transmembrane segments of the protein (TM6), the third segment (p.R144G) and the fourth segment (p.R198Q and p.R213W). The p.R144G and p.R198Q variants have been reported to exhibit a gain-of-function in heterologous expression studies, as p.R144G associated with BFNS1 and R198Q with EEIE7.[40,42] The *KCNQ2* p.R213W variant has been previously described and related to BFNS1[43] as well as in severely affected individuals.[41] In fact, it is known that *KCNQ2* mutation carriers, who have children affected with a severe epileptic phenotype, are mosaic for these variants and often present with BFNS1 associations. Therefore, it is important to look at the clinical history of the carrier mother and study the possibility of mosaicism in this case.[41]

We had the complete clinical information of patients 29 (p.R198Q), 30 (p.R213W), 31 (p.A609T) and 32 (p.R871G). All of them had the most characteristic features of RTT: a normal development during the first months of life, followed by a profound mental retardation, with developmental regression and autistic features. Specifically, patient 29 presented with similar clinical features to the patients previously described with the same variant.[42] Patients 29, 30 and 32 presented with generalized myoclonic or tonic-clonic seizures before the first 18 months of life. In previously described patients, the seizures resolved with treatment (ketogenic diet, carbamazepine or levetiracetam), suggesting that our patients, who are younger than those previously described, may become seizure-free with treatment.[44]

3.5. MEF2C *gene*

MEF2C haploinsufficiency syndrome has been recognized as a neurodevelopmental disorder. To date, fourteen patients with *MEF2C* variants have been identified, including three nonsense, three missense and three frameshift variants.[45–47] We have detected one missense, one no-stop and two frameshift variants. The missense variant is located in the MADS domain; to date, three other missense variants have been described. Although we did not perform functional studies, Zweier et al. (2010) demonstrated that mutations in this domain affect DNA binding specificity; thus, we hypothesize that the same mechanism may contribute to this case.[45] Both frameshift variants are predicted to generate a premature stop codon at amino acid positions 188 and 327, which would result in the loss of the functional TADII domain and an aberrant protein structure. Although the no-stop variant is *de novo* and results in a prolonged protein with 58 extra amino acids, it is predicted to be benign, and, in the gnomAD database, there is one no-stop variant reported (1/244656), p.*484Argext*57. When reviewing this single case, it appears to be a mosaicism, as the wild type is represented 170 times and the alternate variant 54 times. Transcriptional reporter assays have also indicated that *MEF2C* mutations diminish the synergistic transactivation of E-box promoters, including those from *MECP2* and *CDKL5*.[45]

By comparing the clinical descriptions of our patients, most of them have similar facial phenotypes (patients 33, 34 and 35, Supplementary Fig. S1) to those previously described.[46,47] Our patients also presented with epilepsy with no refractoriness that was controlled by anti-epileptic drugs (AEDs), except patient 36 (p.*473Lnext*58), who is seizure-free. Patients 34 and 36 had achieved independent walking, though patients 33 and 35 were able to walk with support despite their unstable wide-based gait. Patient 33 had been followed for a long period and had not improved and had not lost her ability to walk in the last 18 years. Hand stereotypies were present in all our patients and had not diminished during follow–up, as is seen in classic RTT.

Of the four patients who presented with severe intellectual disability with autistic features, three of them displayed content behaviours, which have been described in other patients.[47] Patients 33 and 34 were able to say only a few words, and patients 35 and 36 did not speak at all; these observations are not related to the mutation types observed in other studies, where microdeletions are not always associated with epilepsy.[47,48] Two of our patients had sleep problems (patients 33 and 35) and required pharmacological treatment.

The follow-up of our patients until adulthood will add new details to their phenotype. Our results provide more evidence to support the involvement of *MEF2C* in an RTT-like neurodevelopmental disorder that is characterized by severe intellectual disability, absent or delayed speech, motor and behavioural alterations, early onset of seizures, variable MRI abnormalities and facial dimorphisms.

3.6. SYNGAP1 *gene*

We found four RTT-like patients with four different variants in the *SYNGAP1* gene (synaptic RAS-GTPase-activating protein 1, OMIM#603384), one missense, one nonsense and two frameshift variants. The wild-type protein is localized to dendritic spines in neocortical pyramidal neurons, where it is able to positively or negatively regulate the density of NMDA and AMPA receptors at glutamatergic synapses and mediate signalling downstream of glutamate receptor activation.[49] Severe *de novo* variants in *SYNGAP1* resulting in haploinsufficiency lead to a defined phenotype characterized by ID with epilepsy [termed Mental Retardation-Type 5 (MRD5), OMIM#306684].[50]

The two frameshift variants (p.L595Cfs*55 and p.T674Pfs*36) are clearly deleterious, as they abolish functional domains and are located upstream of other truncating mutations previously reported in ID patients.[51] p.R967* is located in the C-terminus of the protein, which could go through NMD (Nonsense Mediated Decay), which is a missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease.[51] Although p.R905C is predicted by *in silico* studies to be deleterious, its pathogenicity is difficult to assess. On the one hand, this variant is not located in any major domain of the protein; however, on the other hand, it has not been reported in any control databases. It is essential to determinate the inheritance of this variant to establish its causality.

We had complete clinical information from two of the four patients (p.L595Cfs*55 and p.R905C). In line with previous series, hypotonia and gait abnormalities were the main recurrent features. Corroborating previous results,[52] MRI scan in our patients did not show any specific features; thus, brain imaging is not helpful in the diagnoses of *SYNGAP1*-related disorders. Interestingly, our study is the first to report a group of patients with a moderate mental retardation; all of them are able to walk and communicate with few words.

4. Conclusion

In our findings, the limited number of patients grouped by altered genes do not allow to generate a clear genotype-phenotype correlation among them. We only found that RTT-like patients with *SYNGAP1* variants presented the most moderate phenotype, which consists of the possibility of learning, in contrast with the other RTT-like patients. A possible explanation for the heterogeneity in disease onset and/or severity of disease progression is the altered expression of a modifier gene or genes that can exacerbate or diminish the clinical syndrome.[34] Moreover, microenvironmental factors might account for skewing the genotype-phenotype relationship.

NGS methodologies have overcome some difficulties, and the process of diagnosis has moved from one of clinical assessment to one of genetic confirmation. Certainly, there are benefits of defining patients by clinical descriptions, such as grouping patients with similar features for the purpose of clinical management. For RTT molecular diagnosis, it is important to enlarge the study from the three classic genes (*MECP2*, *CDKL5* and *FOXG1*) to include those genes that have a clinical presentation that overlaps with RTT features, such as *STXBP1*, *TCF4*, *SCN2A*, *KCNQ2*, *MEF2C* and *SYNGAP1*. The detection of variants in RTT-like genes may modify the initial clinical diagnosis to other neurodevelopmental syndromes or determine new candidate genes related to RTT-like features, providing the clinician with more information and clues that could help in the prevention of future symptoms or in pharmacologic therapy.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We thank all patients and their families who contributed to this study. The work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01159), the crowd-funding program PRECIPITA, the Spanish Ministry of Health (Fundación Española para la Ciencia y la Tecnología, Spain), the Catalan Association for Rett Syndrome, Fondobiorett and Mi Princesa Rett.

Appendix A

Working group

Hospital Sant Joan de Déu (Barcelona)
Maria del Mar O'Callaghan mocallaghan@sjdhospitalbarcelona.org
Àngels Garcia-Cazorla agarcia@sjdhospitalbarcelona.org
Hospital Sant Joan de Déu (Martorell)
Maria del Carmen Serrano Munuera cserrano@hmartorell.es
Silvia Cuso García scuso@hmartorell.es
Hospital San Borja Arriaran, Universidad de Chile Santiago (Chile)
Monica Troncoso monicatroncososch@gmail.com
Guillermo Fariña guillermofarina@gmail.com
Hospital Infantil Universitario Niño Jesús (Madrid)
Juan José García Peñas jgarciadelarape.1961@gmail.com
Hospital Universitario de Getafe (Madrid)
Belen Gil Fournier bgil.hugf@salud.madrid.org
Soraya Ramiro León soraya.ramiroleon@salud.madrid.org
Hospital Univesitari Parc Taulí (Sabadell)
Miriam Guitart MGuitart@tauli.cat
Neus Baena NBaena@tauli.cat
Hospital Universitario Araba-Txagorritxu (Vitoria)
Guiomar Perez de Nanclares gnanclares@osakidetza.eus
Intzane Ocio Ocio intzane.ocioocio@osakidetza.eus
Neurología, servicio de Medicina Interna. Hospital Universitario de Fuenlabrada (Madrid)
Eva Gutiérrez-Delicado eva.gutierrez@salud.madrid.org
Belén Abarrategui belen.abarrategui@salud.madrid.org
Instituto de Genética Médica y Molecular (INGEMM), IdiPAZ (Madrid)
Eva Barroso eva.barroso@salud.madrid.org
Fernando Santos-Simarro fsantossimarro@gmail.com
Pablo Lapunzina pablo.lapunzina@salud.madrid.org
Análisis Clínicos, Hospital Universitario Fundación Alcorcón (Madrid)
Francisco J. García fjgarciai@fhalcorcon.es
Juan M. Acedo jmacedo@halcorcon.es
Unidad Neuropediátrica, Hospital Universitario Fundación Alcorcón (Madrid)
Asunción García agarcia@fhalcorcon.es
Miguel A. Martinez MAMartinezg@fhalcorcon.es
Servicio de Neurología, Hospital Infantil, Hospital Universitario La Paz (Madrid)
Antonio Martínez-Bermejo ambermejo@salud.madrid.org.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejpn.2019.04.006.

REFERENCES

1. Laurvick CL, de Klerk N, Bower C, Christodoulou J, Ravine D, Ellaway C, et al. Rett syndrome in Australia: a review of the epidemiology. *J Pediatr* 2006;**148**(3):347–52.
2. Rett A. On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wien Med Wochenschr* 1966;**116**(37):723–6. Sep 10.
3. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol* 1983;**14**(4):471–9.
4. Allanson JE, Hennekam RCM, Moog U, Smeets EE. Rett syndrome: a study of the face. *Am J Med Genet Part A* 2011;**155**(7):1563–7.
5. Neul JL, Lane JB, Lee H-S, Geerts S, Barrish JO, Annese F, et al. Developmental delay in Rett syndrome: data from the natural history study. *J Neurodev Disord [Internet]* 2014;**6**(1):20. Available from: http://jneurodevdisorders.biomedcentral.com/articles/10.1186/1866-1955-6-20.
6. Neul JL, Kaufmann WE, Glaze DG, Christodoulou J, Clarke AJ, Bahi-Buisson N, et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol* 2010;**68**(6):944–50.
7. Armani R, Archer H, Clarke A, Vasudevan P, Zweier C, Ho G, et al. Transcription Factor 4 and Myocyte Enhancer Factor 2C mutations are not common causes of Rett syndrome. *Am J Med Genet Part A* 2012;**158**(4):713–9.
8. Lucariello M, Vidal E, Vidal S, Saez M, Roa L, Huertas D, et al. Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype. *Hum Genet* 2016;**135**(12):1343–54.
9. Vidal S, Brandi N, Pacheco P, Gerotina E, Blasco L, Trotta JR, et al. The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome. *Sci Rep* 2017;**7**(1).
10. Srivastava S, Desai S, Cohen J, Smith-Hicks C, Barañano K, Fatemi A, et al. Monogenic disorders that mimic the phenotype of Rett syndrome. *Neurogenetics* 2018;**19**(1):41–7.
11. Schönewolf-Greulich B, Bisgaard AM, Møller RS, Dunø M, Brøndum-Nielsen K, Kaur S, et al. Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin Genet* 2019 Feb;**95**(2):221–30.
12. Carlson RV, Boyd KM, Webb DJ. The revision of the Declaration of Helsinki: past, present and future. *Br J Clin Pharmacol* 2004;**57**:695–713.
13. Monrós E, Armstrong J, Aibar E, Poo P, Canós I, Pineda M. Rett syndrome in Spain: mutation analysis and clinical correlations. *Brain Dev* 2001;**23**(Suppl 1):S251–3.
14. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J [Internet]* 2011;**17**(1):10. Available from: http://journal.embnet.org/index.php/embnetjournal/article/view/200.
15. Andrews S. *FastQC: a quality control tool for high throughput sequence data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. 2013.
16. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;**25**(14):1754–60.
17. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**(6):841–2.
18. Broad Institute. *Picard tools [Internet]*. 2016. Available from: https://broadinstitute.github.io/picard/.
19. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;**27**(21):2987–93.
20. Bateson ZW, Hammerly SC, Johnson JA, Morrow ME, Whittingham LA, Dunn PO. Specific alleles at immune genes, rather than genome-wide heterozygosity, are related to immunity and survival in the critically endangered Attwater's prairie-chicken [cited 2018 Jul 11] *Mol Ecol [Internet]* 2016 Jul 17;**25**(19):4730–44. Available from: http://arxiv.org/abs/1207.3907.
21. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;**22**(3):568–76.
22. Depristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;**43**(5):491–501.
23. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 2012;**28**(21):2747–54.
24. Swanson DA, Steel JM, Valle D. Identification and characterization of the human ortholog of rat STXBP1, a protein implicated in vesicle trafficking and neurotransmitter release. *Genomics* 1998;**48**(3):373–6.
25. Toonen RFG, Wierda K, Sons MS, de Wit H, Cornelisse LN, Brussaard A, et al. Munc18-1 expression levels control synapse recovery by regulating readily releasable pool size. *Proc Natl Acad Sci India* 2006;**103**(48):18332–7. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0608507103.
26. Medrihan L, Tantalaki E, Aramuni G, Sargsyan V, Dudanova I, Missler M, et al. Early defects of GABAergic synapses in the brain stem of a MeCP2 mouse model of Rett syndrome. *J Neurophysiol* 2008;**99**(1):112–21.
27. Olson HE, Tambunan D, Lacoursiere C, Goldenberg M, Pinsky R, Martin E, et al. Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. *Am J Med Genet Part A* 2015;**167**(9):2017–25.
28. Geisheker MR, Heymann G, Wang T, Coe BP, Turner TN, Stessman HAF, et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat Neurosci* 2017;**20**(8):1043–51.
29. Stamberger H, Nikanorova M, Willemsen MH, Accorsi P, Angriman M, Baier H, et al. STXBP1 encephalopathy. *Neurology* 2016;**86**(10):954–62.
30. Sepp M, Pruunsild P, Timmusk T. Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum Mol Genet* 2012;**21**(13):2873–88.
31. Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, et al. Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am J Hum Genet [Internet]* 2007;**80**(5):994–1001. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0002929707609562.
32. De Winter CF, Baas M, Bijlsma EK, Van Heukelingen J, Routledge S, Hennekam RCM. Phenotype and natural history in 101 individuals with Pitt-Hopkins syndrome through an internet questionnaire system. *Orphanet J Rare Dis* 2016;**11**(1).
33. Rosenfeld JA, Leppig K, Ballif BC, Thiese H, Erdie-Lalena C, Bawle E, et al. Genotype-phenotype analysis of TCF4 mutations causing Pitt-Hopkins syndrome shows increased seizure activity with missense mutations. *Genet Med* 2009;**11**(11):797–805.
34. Sanders SJ, Campbell AJ, Cottrell JR, Moller RS, Wagner FF, Auldridge AL, et al. Progress in understanding and treating SCN2A-mediated disorders. *Trends Neurosci* 2018;**41**(7):442–56.
35. Nakamura K, Kato M, Osaka H, Yamashita S, Nakagawa E, Haginoya K, et al. Clinical spectrum of SCN2A mutations

expanding to Ohtahara syndrome. *Neurology* 2013;**81**(11):992–8.
36. Wolff M, Johannesen KM, Hedrich UBS, Masnada S, Rubboli G, Gardella E, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain* 2017;**140**(5):1316–36.
37. Møller RS, Larsen LHG, Johannesen KM, Talvik I, Talvik T, Vaher U, et al. Gene panel testing in epileptic encephalopathies and familial epilepsies. *Mol Syndromol* 2016;**7**(4):210–9.
38. Kong Y, Yan K, Hu L, Wang M, Dong X, Lu Y, et al. Association between SCN1A and SCN2A mutations and clinical/EEG features in Chinese patients from epilepsy or severe seizures. *Clin Chim Acta* 2018;**483**:14–9.
39. Brown DA, Passmore GM. Neural KCNQ (Kv7) channels. *Br J Pharmacol* 2009;**156**:1185–95.
40. Miceli F, Soldovieri MV, Ambrosino P, De Maria M, Migliore M, Migliore R, et al. Early-onset epileptic encephalopathy caused by gain-of-function mutations in the voltage sensor of Kv7.2 and Kv7.3 potassium channel subunits. *J Neurosci [Internet]* 2015;**35**(9):3782–93. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.4423-14.2015.
41. Milh M, Lacoste C, Cacciagli P, Abidi A, Sutera-Sardo J, Tzelepis I, et al. Variable clinical expression in patients with mosaicism for KCNQ2 mutations. *Am J Med Genet Part A* 2015;**167**(10):2314–8.
42. Millichap JJ, Miceli F, De Maria M, Keator C, Joshi N, Tran B, et al. Infantile spasms and encephalopathy without preceding neonatal seizures caused by KCNQ2 R198Q, a gain-of-function variant. *Epilepsia* 2017;**58**(1):e10–5.
43. Sadewa AH, Sasongko TH, Gunadi, Lee MJ, Daikoku K, Yamamoto A, et al. Germ-line mutation of KCNQ2, p.R213W, in a Japanese family with benign familial neonatal convulsion. *Pediatr Int* 2008;**50**(2):167–71.
44. Kato M, Yamagata T, Kubota M, Arai H, Yamashita S, Nakagawa T, et al. Clinical spectrum of early onset epileptic encephalopathies caused by KCNQ2 mutation. *Epilepsia* 2013;**54**(7):1282–7.
45. Zweier M, Gregor A, Zweier C, Engels H, Sticht H, Wohlleber E, et al. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum Mutat* 2010;**31**(6):722–33.
46. Bienvenu T, Diebold B, Chelly J, Isidor B. Refining the phenotype associated with MEF2C point mutations. *Neurogenetics* 2013;**14**(1):71–5.
47. Rocha H, Sampaio M, Rocha R, Fernandes S, Leão M. MEF2C haploinsufficiency syndrome: report of a new MEF2C mutation and review. *Eur J Med Genet* 2016;**59**:478–82.
48. Vrečar I, Innes J, Jones E, Kingston H, Reardon W, Kerr B, et al. Further clinical delineation of the MEF2C haploinsufficiency syndrome: report on new cases and literature review of severe neurodevelopmental disorders presenting with seizures, absent speech, and involuntary movements. *J Pediatr Genet [Internet]* 2017;**06**(03):129–41. Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-0037-1601335.
49. Berryer MH, Hamdan FF, Klitten LL, Møller RS, Carmant L, Schwartzentruber J, et al. Mutations in SYNGAP1 cause intellectual disability, autism, and a specific form of epilepsy by inducing haploinsufficiency. *Hum Mutat* 2013;**34**(2):385–94.
50. Kilinc M, Creson T, Rojas C, Aceti M, Ellegood J, Vaissiere T, et al. Species-conserved SYNGAP1 phenotypes associated with neurodevelopmental disorders. *Mol Cell Neurosci* 2018;**91**:140–50.
51. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 2014;**515**(7526):209–15.
52. Mignot C, von Stülpnage C, Nava C, Ville D, Sanlaville D, Lesca G, et al. Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet* 2016;**53**(8):511–22.

Publicación 5

Título: *Genetic Landscape of Rett Syndrome Spectrum: Improvements and Challenges.*

Autores: Silvia Vidal, Clara Xiol, Ainhoa Pascual-Alonso, del Mar O'Callaghan, Mercè Pineda and Judith Armstrong.

Referencia: *International Journal of Molecular Sciences*. 2019 Aug 12;20(16).

Resumen:

En esta revisión hacemos un repaso de los estudios moleculares actuales, que investigan las causas genéticas del RTT o fenotipos similares que se superponen con otros trastornos genéticos. Esta revisión también subraya la heterogeneidad clínica y genética del RTT y proporciona una visión general de los genes relacionados descritos hasta la fecha, muchos de los cuales están involucrados en la regulación de genes epigenéticos, la acción de neurotransmisores o la transcripción / traducción de ARN. Finalmente, queremos remarcar la importancia de incluir tanto el diagnóstico fenotípico como el genético para proporcionar un asesoramiento genético adecuado desde la perspectiva del paciente y el tratamiento adecuado.

International Journal of *Molecular Sciences*

MDPI

Review

Genetic Landscape of Rett Syndrome Spectrum: Improvements and Challenges

Silvia Vidal [1,2], Clara Xiol [1,2], Ainhoa Pascual-Alonso [1,2], M. O'Callaghan [2,3,4], Mercè Pineda [1] and Judith Armstrong [2,4,5,*]

1 Sant Joan de Déu Research Foundation, 08950 Barcelona, Spain
2 Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, 08950 Barcelona, Spain
3 Neurology Service, Hospital Sant Joan de Déu, 08950 Barcelona, Spain
4 CIBER-ER (Biomedical Network Research Center for Rare Diseases), Institute of Health Carlos III (ISCIII), 28029 Madrid, Spain
5 Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, 08950 Barcelona, Spain
* Correspondence: jarmstrong@sjdhospitalbarcelona.org; Tel.: +34-93-600-9451; Fax: +34-93-600-9760

Received: 3 July 2019; Accepted: 10 August 2019; Published: 12 August 2019

Abstract: Rett syndrome (RTT) is an early-onset neurodevelopmental disorder that primarily affects females, resulting in severe cognitive and physical disabilities, and is one of the most prevalent causes of intellectual disability in females. More than fifty years after the first publication on Rett syndrome, and almost two decades since the first report linking RTT to the *MECP2* gene, the research community's effort is focused on obtaining a better understanding of the genetics and the complex biology of RTT and Rett-like phenotypes without *MECP2* mutations. Herein, we review the current molecular genetic studies, which investigate the genetic causes of RTT or Rett-like phenotypes which overlap with other genetic disorders and document the swift evolution of the techniques and methodologies employed. This review also underlines the clinical and genetic heterogeneity of the Rett syndrome spectrum and provides an overview of the RTT-related genes described to date, many of which are involved in epigenetic gene regulation, neurotransmitter action or RNA transcription/translation. Finally, it discusses the importance of including both phenotypic and genetic diagnosis to provide proper genetic counselling from a patient's perspective and the appropriate treatment.

Keywords: Rett syndrome; Rett-like; NGS; genetics

1. Introduction

Rett syndrome (OMIM#312750) is an early-onset neurodevelopmental disorder, which was first described by Doctor Andreas Rett in 1966 [1]. However, it was not until 1999 when Zoghbi's laboratory identified mutations in the X-linked methyl-CpG-binding protein 2 gene (*MECP2*; OMIM*300005) in RTT patients. This gene encodes a chromatin-associated protein that contains a methyl-CpG binding domain and can activate and repress transcription; it is essential for the maturation of neurons and normal function of nerves cells [2]. The molecular pathogenesis of *MECP2* mutations is complex, involving multiple functions and tissues. MeCP2 has two differentially spliced isoforms of exons 1 and 2 (MeCP2-e1 and MeCP2-e2) which contribute to the diverse functions of MeCP2, but only mutations in exon 1, not exon 2, are observed in RTT. However, the majority of RTT mutations occur in *MECP2* exons 3 and 4 [3]. MeCP2-e1 contains three exons and the start codon is located in exon 1, while MeCP2-e2 contains exons 2, 3 and 4 and the start codon is located in exon 2 [4]. Moreover, MeCP2-e1 is mainly expressed in the central nervous system [3], suggesting that it is the dominant isoform in the brain [5]. Studies using mice models have shown that *Mecp2-e1* deficiency alone contributes to neurologic symptoms, while *Mecp2-e2* has no essential function in the nervous

system [5]. More than 800 different mutations in *MECP2* have been identified in more than 95% of patients with classic RTT and 75% of patients with atypical RTT (RettBASE: MECP2 Variation Database: http://mecp2.chw.edu.au/) [6,7]. There are also some atypical RTT variants, such as the early onset seizure variant and the congenital variant, which have been associated with mutations in cyclin-dependent kinase-like 5 (CDKL5; Xp22; OMIM*300203) and forkhead box protein G1 (FOXG1; 14q12; OMIM*164874), respectively [8,9]. However, the etiology of a subset of patients with a clinical diagnosis of RTT still remains unknown.

Nowadays, with the increasing use of Next-Generation Sequencing (NGS) techniques, and the improvement of the techniques themselves and bioinformatics analysis tools, more patients can obtain a genetic diagnosis, which is important for proper genetic counselling, the patient's future perspective, and treatment options. Consequently, the number of known genes which are disease-causing for RTT-like phenotypes increased remarkably in the last years. This development can actually be observed in the hugely heterogeneous group of neurodevelopmental disorders [10]. This study underlines the current molecular genetic studies performed in RTT patients, highlights the phenotype overlap with other monogenic disorders, and reviews the new treatments that are being performed.

2. RTT and RTT-Like Syndrome

The diagnosis criteria used to establish the clinical diagnosis of RTT was described in 2002 and revised in 2010 [11]. The patients diagnosed with classical RTT should present four main criteria: Partial or complete loss of spoken language, partial or complete loss of purposeful hand movements, gait abnormality, and stereotypic hand movements. In contrast to classical RTT, a diagnosis of atypical RTT is described when the patient present two or more of the main criteria in addition to five or more of the supportive criteria. Nearly all classical RTT patients are characterized by a period of apparently normal development followed by a regression phase, but some atypical forms are congenital and early seizure shows developmental impairment/delay from the first months of life. In addition, there are some clinical features that can exclude a diagnosis of classical RTT, such as brain injury, a neurometabolic disease, or neurological infection [12]. Nowadays, in RTT, as in other diseases, the term "like" is used in patients that do not fulfill established clinical criteria, but present an overlapping phenotype with the disease. Formal consensus criteria for a Rett-like syndrome (RTT-like) are not published yet and a combination of distinct features of RTT can be described as RTT-like phenotype.

The report of pathogenic or likely pathogenic variants in different genes in patients with overlapping phenotypes creates a huge challenge in the clinical diagnosis. NGS, such as gene panels or whole exome/genome sequencing, allows us to solve difficulties and improves results, complementing the clinical diagnosis with a genetics diagnosis. However, making an accurate phenotypic description of the patients is crucial to enable the selection of the most relevant genes to be analyzed and for assessing the clinical significance of genomic variants identified in them.

3. New Technologies for a Rare Genetic Diagnosis

In recent years, NGS technology—a method of simultaneously sequencing millions of fragments of DNA—has emerged as a powerful tool for the study of this type of genetic disease. Now, with the possibility of multiplexing genes and patients, sequencing them at the same time, the cost-efficiency of the technique is comparable to the Sanger sequencing analysis of a single gene [13]. Therefore, the global implementation of these technologies in research laboratories has led to an important increase in the identification of diseases or genes related to RTT/RTT-like phenotype that in some cases had previously been associated with other well-described diseases [14–16]. While the added value of NGS diagnostics in all of these patients is clear, an optimal implementation strategy for diagnostic laboratories is yet to be established [17].

Basically, there are three NGS approaches for DNA sequencing which can be used to improve the diagnostic rate in this hugely heterogeneous group of diseases: (1) Targeted enrichment of a set of genes (gene panels); (2) whole-exome sequencing (WES); (3) whole-genome sequencing (WGS).

Targeted panels focus on individual genes, specific regions of interest, or a subset of genes associated with a wide variety of inherited disorders. This approach is usually the first line of testing, while WES is reserved for cases in which targeted testing has been uninformative [18]. Moreover, panels can be customized and optimized for different regions and sample types, allowing determination of single nucleotide variants (SNVs) from NGS in a more cost-effective manner. The targeted panels are constantly improving because with basic research and WES and WGS of patients without a genetic diagnosis, new genes are discovered or their functions are more clearly understood and, subsequently, are associated with human diseases. For this reason, targeted panels are the best approach in terms of genetics diagnosis.

WES testing often involves testing the child and both parents (trio testing) to assist in the interpretation of variants [19,20]. The current challenge of WES is to determine benign variants from pathological variants, since the exome of a healthy person reveals about 30,000–100,000 variations if compared with the standard reference genome. Using variation databases and software tools, such as the ones that are described in Table 1, the potential disease-causing variation can be detected, although the exact method is uncertain and functional studies continue to be necessary to fully demonstrate the pathogenicity of the variations found [21]. However, mutations in intronic and promoter regions are not covered and nor are greater structural variants like inversions and translocations.

Table 1. List of databases and software tools used in variant analysis.

Data Bases	Description	Website
Human mutation database (HGMD)	Database that represents an attempt to collate all known (published) gene lesions responsible for human inherited disease.	www.hgmd.cf.ac.uk/
Varsome	The human genomic variant search engine.	https://varsome.com/
GnomAD	Data from exome and genome sequencing from a variety of large-scale sequencing projects.	https://gnomad.broadinstitute.org/
dbSNP	Public-domain archive for a broad collection of simple genetic polymorphisms.	www.ncbi.nlm.nih.gov/snp/
ClinVar	Public archive of reports of the relationships among human variations and phenotypes, with supporting evidence.	www.ncbi.nlm.nih.gov/clinvar/
Specific disease databases	Databases such as RettBASE that are freely-available resources for mutation and polymorphism data pertaining to Rett syndrome and other related clinical disorders.	mecp2.chw.edu.au
Software Tools	**Description**	**Website**
Mutation Taster	An in silico prediction tool for the pathogenicity of a variant based on evolutionary conservation, splice-site, mRNA, protein and regulatory features.	www.mutationtaster.org/
SIFT	An in silico prediction tool for nonsynonymous variants based on sequence homology derived from closely related sequences collected through PSI-BLAST.	https://sift.bii.a-star.edu.sg/
Polyphen-2	Tool which predicts possible impact of an amino acid substitution on the structure and function of a human protein using straightforward physical and comparative considerations.	genetics.bwh.harvard.edu/pph2/
Provean	An in silico tool that predicts how nonsynonymous or in-frame indel variant will affect a protein's biological function.	provean.jcvi.org/
Humans Splicing Finder	This tool is aimed to help studying the pre-mRNA splicing.	http://www.umd.be/HSF/

Nowadays, WGS is considered to be the most comprehensive genetic test available, but it is not applied to patient diagnostics because of the complex and challenging data analysis, the high cost compared to targeted panels and WES, and the unknown diagnosis potential of the test. With the sequencing of the whole genome we can detect, besides from SNV in coding regions such as WES, variations in non-coding regions, but in the majority of cases functional studies are required to determine their pathogenicity. To date, there are no publications about WGS in a cohort of RTT nor RTT-like, but studies in intellectual disability (ID) or other diseases are published [22,23]. However, targeted panels have the best cost-efficiency value.

Moreover, NGS not only allows for the detection of SNV variants but with this technology we are also able to detect copy number variations (CNVs); the size and complexity of the genomic regions of interest will determine which NGS method is the best to use. WGS offers the potential to capture all genetic variations, including CNVs and structural changes such as inversions and translocations. WES can detect indels (CNVs ~1–100 bp), but approaches are steadily improving to provide data suitable for larger CNVs because WES covers many different regions and it can be particularly tricky to optimize for uniformity of coverage. Finally, targeted NGS panels offer high uniformity of coverage of targeted regions, and the targeted nature results in lower costs with increasing depth, opening up the possibility for reliable CNV calling.

Recent studies have shown that germline and somatic mosaicism is present in genes related with severe encephalopathies such as Dravet syndrome [24,25], focal cortical dysplasia [26] and intellectual disability [27]. Mosaicism is being postulated as the cause to explain differential phenotype expression of the disease among patients (somatic mosaicism due to a postzygotic mutation) and to explain recurrent mutations in the same family assumed de novo (due to low-grade parental mosaicism). To study mosaicism, other techniques besides NGS must be considered, such as single-molecule molecular inversion probes (smMIP), a technique with high sensitivity for detecting low-grade mosaic variants. Using this technique, it has been shown that parental mosaicism occurs in a substantial proportion of families with mutations in *SCN1A* (7%–13%), which has important implications for recurrence risks in subsequent pregnancies. On the other hand, to study disease severity ranges between patients with mutations in the same gene, a better prediction technique is needed because Dravet syndrome patients with mosaicism have milder phenotypes than those with heterozygous mutations [28].

4. NGS Results: Many Genes, Many Disorders

Apart from RTT and RTT-like patients, which have mutations in the *MEPC2*, *CDKL5* and *FOXG1* genes, there are a percentage of patients without a genetic diagnosis. Now, with WES and panels that incorporate more and more genes related to the central nervous system, the number of patients with a pathogenic variant detected has increased in genes that were previously not related to RTT nor RTT-like. All studies published about these genes are summarized in Table 2.

Using NGS, in only five years more than eighty genes were related to the RTT-like phenotype and some of these genes were identified as causative for aRTT or RTT-like phenotype in these patients, although some of them were associated with well-known syndromes such as Pitt–Hopkins syndrome (*TCF4*, *CNTNAP2* and *NRXN1* genes), Phelan–McDermid syndrome (*SHANK3* gene), Angelman syndrome (*UBE3A* gene), Kleefstra syndrome (*EHMT1*) and Cornelia de Lange syndrome (*SMC1A*). In addition, a substantial number of genes are epileptic encephalopathy genes (*STXBP1*, *SCN1A*, *SCN2A*, *SCN8A*, *GRIN2A*, *GRIN2B*, *HCN1*, *SLC6A1*, *KCNA2*, *EEF1A2*, *KCNB1* and *SYNGAP1*) and are related with mental retardation and epilepsy (*IQSEC2* and *MEF2C*).

Note that, regarding the number of patients described in these publications, there are some genes that are more represented (Figure 1). *STXBP1*, *TCF4*, *SCN2A*, *WDR45*, and *MEF2C* are the most common genes with a pathogenic (or likely pathogenic) variant detected in patients with RTT/RTT-like phenotype.

Table 2. List of recent publications about genetic studies in Rett syndrome (RTT) and the genes reported.

Publications	Genes
Gilissen et al. 2014 [23]	*SMC1A*
Baasch et al. 2014 [29]	*CN2A*
Saitsu et al. 2014 [30]	*TBL1XR1*
Okamoto et al. 2015 [31]	*GABRD*
Hara et al. 2015 [32]	*SHANK3*
Olson et al. 2015 [33]	*STXBP1, SCN8A, IQSEC2*
Hoffjan et al. 2016 [34]	*WDR45*
Lee et al. 2016 [35]	*SATB2*
Saez et al. 2016 [36]	*JMJD1C*
Rocha et al. 2016 [37]	*MEF2C*
Lucariello et al. 2016 [19]	*ANKRD31, CHRNA5, HCN1, SCN1A, TCF4, GRIN2B, SLC6A1, MGRN1, BTBD9, SEMA6B, AGAP6, MGRN1,VASH2, ZNF620, GRAMD1A, GABBR2, ATP8B1, HAP1, PDLIM7, SRRM3, CACNA1I*
Lopes et al. 2016 [38]	*TCF4, EEF1A2, STXBP1, ZNF238, SLC35A2, ZFX, SHROOM4, EIF2B2, RHOBTB2, SMARCA1, GABBR2, EIF4G1, HTT*
Vidal et al. 2017 [15]	*GRIN2B, GABBR2, MEF2C, STXBP1, KCNQ2, SLC2A1, TCF4, SCN2A, SYNGAP1, CACNA1I, CHRNA5, HCN1*
Sajan et al. 2017 [39]	*PWP2, SCG2, IZUMO4, XAB2, ZSCAN12, IQSEC2, FAM151A, SYNE2, SMC1A, ARHGEF10L, HDAC1, TAF1B, KCNJ10, CHD4, LRRC40, LAMB2, GRIN2B, IMPDH2, SAFB2, ACTL6B, STXBP1, TRRAP, WDR45, SLC39A13, FAT3, IQGAP3, NCOR2, GABRB2, TCF4, GRIN2A*
Allou et al. 2017 [40]	*IQSEC2, KCNA2*
Yoo et al. 2017 [41]	*GABBR2*
Vuillaume et al. 2018 [42]	*GABBR2*
Huisman et al. 2017 [43]	*SMC1A*
Wang et al. 2018 [14]	*MEF2C*
Percy et al. 2018 [44]	*CTNNB1, WDR45*
Srivastava et al. 2018 [18]	*KCNB1, IQSEC2, MEIS2, TCF4, WDR45*
Iwama et al. 2019 [45]	*ATP6V0A1, USP8, MAST3, NCOR2, WDR45, STXBP1, SHANK3, UBE3A, GABRA1, SCN2A, SCN8A, GRIN2B, IQSEC2, CAMK2B, CUX2, CACNA1D, CACNA1G, ITPR1, KIF1A, SYNGAP1, NALCN, NR2F1, IRF2BPL, MAST1, COL4A1, HDAC8, TCF4, PDHA1, PPT1, DNMT3A, MEF2C*
Schönewolf-Greulich et al. 2019 [16]	*STXBP1, SCN2A, KCNB1, TCF4, SHANK3, SMC1A*
Vidal et al. 2019 [46]	*STXBP1, TCF4, SCN2A, KCNQ2, MEF2C, SYNGAP1*

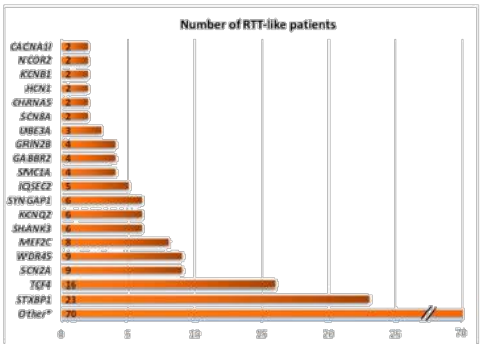


Figure 1. Overall results of all the summarized genetic studies. Patients grouped by gene defect.

The *STXBP1* gene (syntaxin-binding protein 1; OMIM#602926) encodes a transmembrane attachment protein receptor that plays an important role in the release of neurotransmitters via regulation of syntaxin, a transmembrane attachment protein receptor [47]. Pathogenic variants in this gene, reducing its expression, depress the functions of GABAergic and glutamatergic synapses, particularly in GABAergic interneurons [48], a process that has been shown to be altered in RTT patients [49]. Mutations in *STXBP1* have been associated with epileptic encephalopathy, early infantile 4 (EEIE4) and a series of neurodevelopmental disorders, including RTT-like syndrome.

The *TCF4* gene (transcription factor 4; OMIM*602272) is a broadly expressed basic helix–loop–helix (bHLH) protein that functions as a homodimer or as a heterodimer with other bHLH proteins. These dimers bind DNA at Ephrussi (E) box sequences motif ('CANNTG'). Alternative splicing creates numerous N-terminally distinct TCF4 isoforms that differ in their subcellular localization and transactivational capacity [50]. Mutations in *TCF4* have been associated with Pitt–Hopkins syndrome (PTHS), which is characterized by severe ID, delayed motor development, seizures, wide mouth and distinctive facial features, hypotonia, microcephaly, limited walking abilities, and intermittent hyperventilation followed by apnea [31]. The microcephaly, intermittent hyperventilation, and stereotype hand movements may steer clinicians towards a diagnosis of RTT-like rather than PTHS. The distinct facial features presented in patients with a clear PTHS phenotype, such as thin eyebrows, sunken eyes, a pronounced double curve of the upper lip, and a wide mouth with full lips, is more consistent with PTHS and helps to distinguish PTHS from RTT, but some patients do not always present distinctly these facial dismorphisms, or they are often not well-defined during the first year of life [51].

SCN2A (neuronal voltage-gated sodium channel NaV1.2; OMIM*182390) encodes one member of the sodium channel alpha subunit gene family, responsible for generation and propagation of action potentials, chiefly in nerve and muscle. Pathogenic variants in the *SCN2A* gene that produce loss-of-function of the protein lead to ASD/ID and increased channel activity that lead to epileptic encephalopathy early infantile 11 (EEIE11) and benign familial neonatal-infantile seizures (BFIS) [52].

WD40 repeat proteins are an important key component that regulates the assembly of multiprotein complexes by presenting a beta-propeller platform for simultaneous and reversible protein–protein interactions [53]. Variants in *WDR45* (WD repeat-containing protein 45; OMIM*300526) are associated with developmental delay in early childhood and progressive neurodegeneration in adolescence or adulthood related to iron accumulation in the *globus pallidus* and *substantia nigra* [30,54]. Affected patients may have features overlapping those of RTT, including developmental regression, hand-wringing, and seizures. Some may even have a diagnosis of typical or atypical RTT [55].

MEF2C (OMIM*600662) belongs to the myocyte enhancer factor-2 (MEF2) family of transcription factors. *MEF2C* plays an important role in myogenesis, development of the anterior heart field, neural crest and craniofacial development, and neurogenesis, among others [56]. It is well-described the *MEF2C* haploinsufficiency syndrome that has been recognized as a neurodevelopmental disorder. Until now, fourteen patients with point mutation pathogenic, or likely pathogenic, variants in *MEF2C* have been identified in RTT-like patients, including three nonsense, three missense and three frameshift variants [37,57].

In Table 3, the most common genes for RTT-like phenotypes are summarized, and all of them present an autosomal dominant inheritance pattern, either autosomal or linked to the X chromosome and are caused by de novo heterozygous mutations in the germline. Note that all the diseases are severe IDs and most of them meet the four main criteria for RTT; all of them present a loss or speech severe deficit, gait abnormalities and lost or absent purposeful hand movements linked to stereotypical hand movements, such as hand wringing. Only developmental regression, one of the four main criteria, is absent in PTHS. Moreover, other common symptoms in RTT are present in these diseases too, such as epilepsy, breathing disturbances and microcephaly, although not all the RTT patients present them. In contrast, other symptoms that are not present in RTT are present in atypical forms related to *CDKL5* and *FOXG1*, such as dysmorphic facial features in PTHS or CNS abnormalities in neurodegeneration with brain iron accumulation.

Table 3. List of causative genes in RTT-like diagnoses or differential diagnoses and the phenotypical overlap with Rett syndrome (modified from Schönewolf-Greulich et al., 2019 [16]).

			RTT Genes							
			MECP2	***CDKL5***	***FOXG1***	***STXBP1***	***TCF4***	***SCN2A***	***WDR45***	***MEF2C***
		Disorder	Rett syndrome	EEP 2	RTT, congenital variant	EEP 4	Pitt-Hopkins syndrome	EEP 11	Neurodegeneration with brain iron accumulation 5	*MEF2C* haploinsufficiency syndrome
		OMIM#	312750	300672	613454	612164	610954	613721	300894	613443
		Inheritance	XLD	XLD	AD	AD	AD	AD	XLD	AD
Present in RTT	Required	Developmental regression	+	+	+	+	-	+	+	-
	Four main criteria	Purposeful hand movements lost/absent	+	+	+	+	+	+	+	+
		Speech severe deficit/loss	+	+	+	+	+	+	+	+
		Gait abnormality	+	+	+	+	+	+	+	+
		Stereotypic hand movements	+	+	+	+	+	+	+	+
	Other common symptoms	Breathing abnormality	+	+	-	-	+	-	-	-
		ID	+	+	+	+	+	+	+	+
		Epilepsy	+	+	+	+	+	+	+	+
		Microcephaly	+	+	+	+	+	+	+	-
Not present in RTT	Exclusion criteria	CNS abnormality	-	-	+	-	+	-	+	+
	Other symptoms	Dysmorphic facial features	-	+	-	-	+	-	+	+

Abbreviations: EEP, epileptic encephalopathy; XLD, X-linked dominant; AD, autosomal dominant. Plus (+) is noted if the symptom has been described in one or more patients with a pathogenic variant in the gene. The symptoms emphasized are the main clinical features according to the 2010 classification of clinical Rett and other specific features of RTT. The gray colour indicates clinical symptoms in common with RTT.

Focusing on the pathophysiology of RTT at the brain level, female mice heterozygous for the null *MECP2* present microcephaly without gross neuropathological changes. Specifically, mouse and human neurons without MeCP2 have smaller somas and decreased dendritic complexity [58,59]. A decrease in synaptic plasticity and abnormalities in neurotransmitter concentrations is also observed in many neuronal types [60–62]. Looking at the established knockout mice models of those of the most common genes detected in RTT patients, morphological and/or physiological features resemble RTT mouse models. For instance, young adult *Mef2c* and *Scn2a* cKO mice present a normal gross brain morphology and cortical layer organization, as murine RTT models. In contrast, *Mef2c* cKO mice cause an increase in dendritic spine density on dentate granule neurons of the hippocampal dentate gyrus [63] and *Scn2a* cOK in pre-oligodendrocyte alters their morphology, impairs myelination and reduces axon-oligodendrocyte interactions [64,65]. Moreover, depletion of specific presynaptic proteins involved in exocytosis, such as *STXBP1*, causes abnormalities in neurotransmitter concentrations and produces neuronal cell death [66]. *TCF4* haploinsufficiency mice exhibit a delay in neuronal migration, and a significant increase in the number of upper-layer cortical neurons, as well as abnormal dendrite and synapse formation [67]. *WDR45* cOK mice show a loss of neurons in prefrontal cortex and basal ganglion in aged mice and increased apoptosis in prefrontal cortex, recapitulating a hallmark of neurodegeneration [68].

At the moment, more than 80 genes have been associated with RTT/RTT-like, but their link with RTT should be critically evaluated. In order to relate a gene with RTT several points have to be taken into consideration: (1) Proper clinical characterizations of the patients; (2) deep comprehension of the functions of the candidate gene; (3) validated evidence that mutation found it is pathogenic. The more patients with a similar phenotype have pathogenic mutations in the same gene, the more consistent evidence we will have of that gene being part of the list of genes related to RTT.

5. Functions and Pathways around RTT

The list of genes related to RTT/RTT-like phenotypes is complex and diverse. However, using REACTOME (an open-source, open access, manually curated and peer-reviewed pathway database; https://reactome.org) and STRING (a database of known and predicted protein–protein interactions; https://string-db.org/), we can identify some groups of genes with functions involved in common mechanisms. Several pathways can be studied in RTT/RTT-like patients, such as chromatin modulation, synaptic function and ubiquitin conjugation (Figure 2).

This list currently includes 15 genes (*ACTL6B, ANKRD31, CHD4, HDAC1, JMJD1C, MEF2C, NCOR2, SATB2, SMARCA1, TBL1XR1, TRRAP, ZFX. ZNF238, ZNF620* and *ZSCAN12*) involved in chromatin modulation pathways, such as chromatin-modifying enzymes and histone deacetylases (HDACs) and 21 genes (*ATP6V0A1, CACNA1I, CHRNA5, GABBR2, GABRB2, GABRD, GRIN2A, GRIN2B, HCN1, IQSEC2, KCNA2, KCNJ10, KCNQ2, SCG2, SCN1A, SCN2A, SCN8A, SHANK3, SLC6A1, STXBP1* and *SYNGAP1*) involved in synaptic function, necessary for GABAergic, glutamatergic and dopaminergic synapses, synaptic vesicles trafficking, ion homeostasis in neurons and circadian entrainment. The link of all these genes to the same pathways could explain why these patients' phenotypes overlap, causing impaired synaptic function, sleep disturbances and major dysregulation of gene expression. Notably, there are also a few genes (*MGRN1, RHOBTB2* and *USP8*) involved in ubiquitination processes, which *UBE3A* (the gene responsible for Angelman syndrome) is also linked to. The considerable overlapping of RTT and Angelman features could be explained due to this relationship.

MeCP2 performs many tasks during the neurodevelopment, such as regulating the gene expression of other genes, modulating epigenetic imprinting and neurotransmitter actions. Hence it is challenging to create a well-defined pathways involving MeCP2 and predict the downstream effects that disruption of the MeCP2 function can generate. Ehrhart et al. 2016 [69] created a comprehensive visualization of the biologic pathways showing how *MECP2* upstream and downstream regulation developed.

Moreover, it had been and published on WikiPathways which will serve as template for future omics data driven research (http://www.wikipathways.org/instance/WP3584) [69].

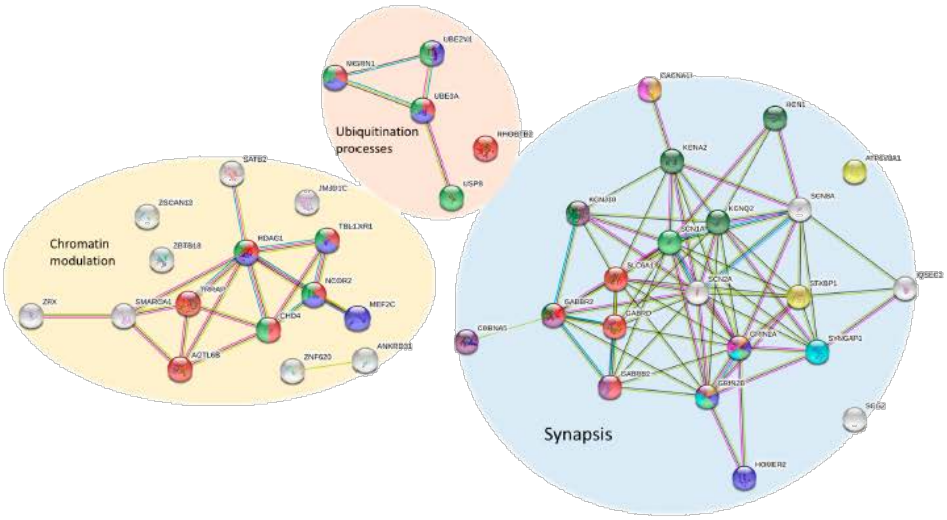


Figure 2. RTT-related protein known functional interaction networks. (A) Chromatin modulation: Red for chromatin modifying enzymes, blue for proteins involved in MECP2-mediated transcriptional regulation and green for histone deacetylases (HDACs). (B) Ubiquitination processes: Red for proteins linked to ubiquitin-mediated proteolysis, green for proteins involved in ubiquitin-like modifier conjugation pathway and blue for ubiquitination and proteasome degradation proteins. (C) Synapsis: Red for proteins involved in GABAergic synapses, dark blue for proteins involved in glutamatergic synapses, light green for proteins involved in dopaminergic synapses, yellow for proteins involved in the synaptic vesicle cycle, pink for proteins in the calcium signalling pathway, dark green for potassium channels, light blue for proteins in the Ras signalling pathway, orange for proteins related to circadian entrainment and purple for neurotransmitter receptors and postsynaptic transmission proteins.

Discovering the pathways related to *MEPC2*, we could better link genes that are mutated in patients without *MECP2* defects. Nowadays, the combination of the omics data analysis and prior knowledge databases are a powerful approach to identify connections between mutation and phenotype. Ehrhart et al. 2019 [70] identified a subset of genes, which are significantly different in several transcriptomics datasets and were not described yet in the context of RTT. They described that these genes are involved in molecular pathways and several processes known to be affected in RTT patients [70]. For example, altered calcium homeostasis seems to be responsible for an abnormal neuronal development and generates epilepsy; and tubulin, *ERM* and *MEF2C* are some of the altered proteins related to cytoskeletal abnormalities that are present not only in RTT, but also in Angelman syndrome [69]. In the same way, cholesterol biosynthesis is altered in RTT and in Smith–Lemli–Opitz syndrome, in which it has been pointed out to be the cause of the autism and malformations [71]. The NF-kB pathway, which is involved in nervous system development, synaptic transmission and cognition, is altered in RTT and RTT-like patients and seems to be the cause of mental retardation [72]. Another pathway observed in patients with RTT, autism, and Parkinson's disease (PD) is the neurotransmitter imbalance of GABA, Glutamine and Dopamine. This imbalance seems to be responsible for the autism features RTT patients present and the motor difficulties that patients with PD and RTT have [73].

An important feature of MeCP2 reduction in RTT mouse models and individuals with RTT is a propensity for seizure, a prominent signature in many brain diseases, including RTT [74,75]. The deletion of MeCP2 from all forebrain GABAergic interneurons recapitulates major phenotype of

RTT [75], demonstrating that altered inhibitory function is critical for normal function of GABA-releasing neurons and an important aspect of RTT pathophysiology. It has been demonstrated recently that human neurons derived from patients with RTT and RTT mouse models show a significant reduced the *SLC12A5* gene expression, resulting in a delayed GABA functional switch [76]. This gene encodes a neuron-specific K+/Cl− cotransporter 2, the major extruder of intracellular chloride in mature neurons. Moreover, it has been established that MeCP2 regulates KCC2 expression through inhibiting RE1-silencing transcriptional factor [77], and it is suggested that KCC2 should play a role in the pathophysiology of RTT

Recently, Cosentino et al. (2019) have described several alterations in RTT patients and animal models during the pre-symptomatic stage [78]. During this stage some compensatory mechanisms keep the phenotypic outcome to a minimum until *MECP2* deficiency cannot be supplied and the known phenotype of RTT becomes apparent [79]. Thus, since the alterations found in RTT spectrum disorders are due to both direct and indirect effects of *MECP2* and related genes' deficiencies/malfunctions, the best approach for characterization and consequent treatment would be the comprehensive study of all the altered pathways that have been discovered in these patients. If we were able to find a way to compensate those altered pathways, a treatment could be implemented in very early stages of development, even before the onset of the most remarkable features of the syndrome [78,79].

6. Future Perspectives and Treatment Options

In recent years, impressive advances have been achieved not only in the genetics diagnosis of neurodevelopmental disorders, but also in elucidating the physiological pathway and molecular mechanisms involved in the clinical manifestations of the diseases. An increasing amount of evidence is emerging that understanding these mechanisms is relevant for the selection of the most appropriate treatment in the affected individual.

More than 50 years have passed since the description and clinical characterization of RTT and the current standard of care for patients remains limited to supportive and symptomatic therapies that can palliate the symptomatology of the patients, but not cure the disease per se. Drug treatment consists mainly of off-label prescriptions due to the lack of approved medications for the disorder. Until now, all RTT trials have been based on the essential role of *MECP2* in the development and maintenance of neurons in the central nervous system, and its specific role in the distinct cellular subtypes focused on the following specific neurotransmitters: dextromethorphan (an NMDA receptor antagonist, mainly used for cognition and seizures [80]); desipramine (a noradrenaline reuptake inhibitor, for breathing abnormalities [81]); and IGF-1 [82]. To date, emulating the MeCP2 function, using a pharmacology strategy, as a treatment for RTT is not the best and most successful strategy.

However, it is not only patients with RTT or RTT-like that are caused by defects in the *MECP2* gene. An improved understanding of the different genes mutated in the RTT-like phenotype generates important therapeutic clues and opportunities to develop novel and better treatments. A treatment which targets neuronal maturational defects seen in *MECP2* mutations may not be effective for an ion channelopathy due to *KCNB1* alterations. In this way, therapies must focus on personalized treatments for each individual, depending on which gene and which type of mutation carries. For example, for the *SCN2A* gene that encodes the voltage-gated sodium channel Nav1.2, gain-of-function versus loss-of-function variants in *SCN2A* determine whether sodium channel blockers improve or worsen seizure control [52]. Another example is that L-serine supplementation might ameliorate *GRIN2B*-related severe encephalopathy [83]. Additionally, for *STXBP1*, which plays an important role in presynaptic vesicle docking and fusion [84], current treatments are largely limited to seizure control and future therapies will also need to target the developmental aspects of the disease [83]. It has been hypothesized that enhancing *KCC2* expression could increase the efficacy of GABAergic inhibition and also improve the dendritic spine and excitatory synapse development, both of which are abnormal in RTT [85,86] and may be caused by an aberrant interaction between *KCC2* and the dendritic cytoskeleton [87,88]. Recently, two works have considered the pathogenic role of diminished

KCC2 expression in the Mecp2 null model: The first found that the use of bumetanide can attenuate the unbalance glutamatergic/GABAergic ratio if treated in the early stages of the disorder [89]; and Tang et al. (2019) have shown that the injection of KEEC KW-2449 or piperine (small-compounds) in Mecp2 mutant mice ameliorated disease-associated respiratory and locomotion phenotypes [90].

7. Conclusions

All of the individuals summarized in this review met the diagnostic criteria for RTT or RTT-like; however, lacking a defect in *MECP2* underscores the importance of carrying out additional genetic testing, whether it is by specific gene panels, WES or WGS, to identify the specific etiology and to direct appropriate diagnostic and therapeutic strategies. Many disorders can be caused by multiple genes, such as West syndrome or Charcot-Marie-Tooth, and these are considered as the same disorder as long as they share a common phenotype. Therefore, we consider that all classical, atypical and all RTT-like phenotypes could be grouped into an RTT spectrum disorder with many causative genes.

From a research point of view, the new genes recently associated with RTT-like phenotypes without a clear description of their biological functions network involved need further study. Furthermore, variants with an unknown clinical significance in genes without clear defined functions have been found through screening RTT spectrum cohorts without genetic diagnosis. These new candidate genes need functional studies to establish their potential role in the disease pathogenesis. From a clinical perspective, a better definition of the pathways that connect of all these genes involved in all RTT and RTT-like phenotypes would enable a better understanding of the genetic landscape of the RTT spectrum. The elucidation of the functional pathways involved in all these patients could support the development of future targeted therapies.

Author Contributions: S.V., C.X., M.P. and J.A. conceived and supervised the review; S.V. and C.X. wrote the original draft. S.V., C.X., A.P.-A., M.O., M.P. and J.A., wrote—reviewed & edited. J.A. and M.P. supervised. All authors reviewed the article critically for intellectual content.

Funding: The work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01159), the crowd-funding program PRECIPITA, the Spanish Ministry of Health (Fundación Española para la Ciencia y la Tecnología), the Catalan Association for Rett Syndrome, FondoBiorett and Mi Princesa Rett.

Conflicts of Interest: The authors declare no competing interests. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Rett, A. On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wien. Med. Wochenschr.* **1966**, *116*, 723–726. [PubMed]
2. Amir, R.E.; Van den Veyver, I.B.; Wan, M.; Tran, C.Q.; Francke, U.; Zoghbi, H.Y. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **1999**, *23*, 185–188. [CrossRef] [PubMed]
3. Kriaucionis, S.; Bird, A. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res.* **2004**, *32*, 1818–1823. [CrossRef] [PubMed]
4. Mnatzakanian, G.N.; Lohi, H.; Munteanu, I.; Alfred, S.E.; Yamada, T.; MacLeod, P.J.; Jones, J.R.; Scherer, S.W.; Schanen, N.C.; Friez, M.J.; et al. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. *Nat. Genet.* **2004**, *36*, 339–341. [CrossRef] [PubMed]
5. Yasui, D.H.; Gonzales, M.L.; Aflatooni, J.O.; Crary, F.K.; Hu, D.J.; Gavino, B.J.; Golub, M.S.; Vincent, J.B.; Carolyn Schanen, N.; Olson, C.O.; et al. Mice with an isoform-ablating Mecp2 exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum. Mol. Genet.* **2014**, *23*, 2447–2458. [CrossRef] [PubMed]
6. Christodoulou, J.; Grimm, A.; Maher, T.; Bennetts, B. RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum. Mutat.* **2003**, *21*, 466–472. [CrossRef] [PubMed]

7. Neul, J.L.; Fang, P.; Barrish, J.; Lane, J.; Caeg, E.B.; Smith, E.O.; Zoghbi, H.; Percy, A.; Glaze, D.G. Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology* **2008**, *70*, 1313–1321. [CrossRef] [PubMed]
8. Kalscheuer, V.M.; Freude, K.; Musante, L.; Jensen, L.R.; Yntema, H.G.; Gecz, J.; Sefiani, A.; Hoffmann, K.; Moser, B.; Haas, S.; et al. Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nat. Genet.* **2003**, *35*, 313–315. [CrossRef]
9. Ariani, F.; Hayek, G.; Rondinella, D.; Artuso, R.; Mencarelli, M.A.; Spanhol-Rosseto, A.; Pollazzon, M.; Buoni, S.; Spiga, O.; Ricciardi, S.; et al. FOXG1 is responsible for the congenital variant of Rett syndrome. *Am. J. Hum. Genet.* **2008**, *83*, 89–93. [CrossRef]
10. Srivastava, S.; Love-Nichols, J.A.; Dies, K.A.; Ledbetter, D.H.; Martin, C.L.; Chung, W.K.; Firth, H.V.; Frazier, T.; Hansen, R.L.; Prock, L.; et al. Meta-analysis and multidisciplinary consensus statement: Exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet. Med.* **2019**. [CrossRef]
11. Hagberg, B.; Hanefeld, F.; Percy, A.; Skjeldal, O. An update on clinically applicable diagnostic criteria in Rett syndrome. Comments to Rett Syndrome Clinical Criteria Consensus Panel Satellite to European Paediatric Neurology Society Meeting, Baden Baden, Germany, 11 September 2001. *Eur. J. Paediatr. Neurol.* **2002**, *6*, 293–297. [CrossRef]
12. Neul, J.L.; Kaufmann, W.E.; Glaze, D.G.; Christodoulou, J.; Clarke, A.J.; Bahi-Buisson, N.; Leonard, H.; Bailey, M.E.; Schanen, N.C.; Zappella, M.; et al. Rett syndrome: Revised diagnostic criteria and nomenclature. *Ann. Neurol.* **2010**, *68*, 944–950. [CrossRef]
13. Koboldt, D.C.; Steinberg, K.M.; Larson, D.E.; Wilson, R.K.; Mardis, E.R. The next-generation sequencing revolution and its impact on genomics. *Cell* **2013**, *155*, 27–38. [CrossRef]
14. Wang, J.; Zhang, Q.; Chen, Y.; Yu, S.; Wu, X.; Bao, X.; Wen, Y. Novel MEF2C point mutations in Chinese patients with Rett (-like) syndrome or non-syndromic intellectual disability: Insights into genotype-phenotype correlation. *BMC Med. Genet.* **2018**, *19*, 191. [CrossRef]
15. Vidal, S.; Brandi, N.; Pacheco, P.; Gerotina, E.; Blasco, L.; Trotta, J.R.; Derdak, S.; Del Mar O'Callaghan, M.; Garcia-Cazorla, A.; Pineda, M.; et al. The utility of Next Generation Sequencing for molecular diagnostics in Rett syndrome. *Sci. Rep.* **2017**, *7*, 12288. [CrossRef]
16. Schonewolf-Greulich, B.; Bisgaard, A.M.; Moller, R.S.; Duno, M.; Brondum-Nielsen, K.; Kaur, S.; Van Bergen, N.J.; Lunke, S.; Eggers, S.; Jespersgaard, C.; et al. Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin. Genet.* **2019**, *95*, 221–230. [CrossRef]
17. Di Resta, C.; Galbiati, S.; Carrera, P.; Ferrari, M. Next-generation sequencing approach for the diagnosis of human diseases: Open challenges and new opportunities. *EJIFCC* **2018**, *29*, 4–14.
18. Srivastava, S.; Desai, S.; Cohen, J.; Smith-Hicks, C.; Baranano, K.; Fatemi, A.; Naidu, S. Monogenic disorders that mimic the phenotype of Rett syndrome. *Neurogenetics* **2018**, *19*, 41–47. [CrossRef]
19. Lucariello, M.; Vidal, E.; Vidal, S.; Saez, M.; Roa, L.; Huertas, D.; Pineda, M.; Dalfo, E.; Dopazo, J.; Jurado, P.; et al. Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype. *Hum. Genet.* **2016**, *135*, 1343–1354. [CrossRef]
20. Mak, C.C.; Leung, G.K.; Mok, G.T.; Yeung, K.S.; Yang, W.; Fung, C.W.; Chan, S.H.; Lee, S.L.; Lee, N.C.; Pfundt, R.; et al. Exome sequencing for paediatric-onset diseases: Impact of the extensive involvement of medical geneticists in the diagnostic odyssey. *NPJ Genom. Med.* **2018**, *3*, 19. [CrossRef]
21. Lelieveld, S.H.; Veltman, J.A.; Gilissen, C. Novel bioinformatic developments for exome sequencing. *Hum. Genet.* **2016**, *135*, 603–614. [CrossRef]
22. Lupski, J.R.; Reid, J.G.; Gonzaga-Jauregui, C.; Rio Deiros, D.; Chen, D.C.; Nazareth, L.; Bainbridge, M.; Dinh, H.; Jing, C.; Wheeler, D.A.; et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N. Engl. J. Med.* **2010**, *362*, 1181–1191. [CrossRef]
23. Gilissen, C.; Hehir-Kwa, J.Y.; Thung, D.T.; van de Vorst, M.; van Bon, B.W.; Willemsen, M.H.; Kwint, M.; Janssen, I.M.; Hoischen, A.; Schenck, A.; et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature* **2014**, *511*, 344–347. [CrossRef]
24. Halvorsen, M.; Petrovski, S.; Shellhaas, R.; Tang, Y.; Crandall, L.; Goldstein, D.; Devinsky, O. Mosaic mutations in early-onset genetic diseases. *Genet. Med.* **2016**, *18*, 746–749. [CrossRef]

25. de Lange, I.M.; Koudijs, M.J.; van 't Slot, R.; Sonsma, A.C.M.; Mulder, F.; Carbo, E.C.; van Kempen, M.J.A.; Nijman, I.J.; Ernst, R.F.; Savelberg, S.M.C.; et al. Assessment of parental mosaicism in SCN1A-related epilepsy by single-molecule molecular inversion probes and next-generation sequencing. *J. Med. Genet.* **2019**, *56*, 75–80. [CrossRef]
26. Uddin, M.; Woodbury-Smith, M.; Chan, A.; Brunga, L.; Lamoureux, S.; Pellecchia, G.; Yuen, R.K.C.; Faheem, M.; Stavropoulos, D.J.; Drake, J.; et al. Germline and somatic mutations in STXBP1 with diverse neurodevelopmental phenotypes. *Neurol. Genet.* **2017**, *3*, e199. [CrossRef]
27. D'Gama, A.M.; Walsh, C.A. Somatic mosaicism and neurodevelopmental disease. *Nat. Neurosci.* **2018**, *21*, 1504–1514. [CrossRef]
28. de Lange, I.M.; Koudijs, M.J.; van 't Slot, R.; Gunning, B.; Sonsma, A.C.M.; van Gemert, L.; Mulder, F.; Carbo, E.C.; van Kempen, M.J.A.; Verbeek, N.E.; et al. Mosaicism of de novo pathogenic SCN1A variants in epilepsy is a frequent phenomenon that correlates with variable phenotypes. *Epilepsia* **2018**, *59*, 690–703. [CrossRef]
29. Baasch, A.L.; Huning, I.; Gilissen, C.; Klepper, J.; Veltman, J.A.; Gillessen-Kaesbach, G.; Hoischen, A.; Lohmann, K. Exome sequencing identifies a de novo SCN2A mutation in a patient with intractable seizures, severe intellectual disability, optic atrophy, muscular hypotonia, and brain abnormalities. *Epilepsia* **2014**, *55*, e25–e29. [CrossRef]
30. Saitsu, H.; Tohyama, J.; Walsh, T.; Kato, M.; Kobayashi, Y.; Lee, M.; Tsurusaki, Y.; Miyake, N.; Goto, Y.; Nishino, I.; et al. A girl with West syndrome and autistic features harboring a de novo TBL1XR1 mutation. *J. Hum. Genet.* **2014**, *59*, 581–583. [CrossRef]
31. Okamoto, N.; Miya, F.; Tsunoda, T.; Kato, M.; Saitoh, S.; Yamasaki, M.; Shimizu, A.; Torii, C.; Kanemura, Y.; Kosaki, K. Targeted next-generation sequencing in the diagnosis of neurodevelopmental disorders. *Clin. Genet.* **2015**, *88*, 288–292. [CrossRef]
32. Hara, M.; Ohba, C.; Yamashita, Y.; Saitsu, H.; Matsumoto, N.; Matsuishi, T. De novo SHANK3 mutation causes Rett syndrome-like phenotype in a female patient. *Am. J. Med. Genet. A* **2015**, *167*, 1593–1596. [CrossRef]
33. Olson, H.E.; Tambunan, D.; LaCoursiere, C.; Goldenberg, M.; Pinsky, R.; Martin, E.; Ho, E.; Khwaja, O.; Kaufmann, W.E.; Poduri, A. Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. *Am. J. Med. Genet. A* **2015**, *167A*, 2017–2025. [CrossRef]
34. Hoffjan, S.; Ibisler, A.; Tschentscher, A.; Dekomien, G.; Bidinost, C.; Rosa, A.L. WDR45 mutations in Rett (-like) syndrome and developmental delay: Case report and an appraisal of the literature. *Mol. Cell Probes* **2016**, *30*, 44–49. [CrossRef]
35. Lee, J.S.; Yoo, Y.; Lim, B.C.; Kim, K.J.; Choi, M.; Chae, J.H. SATB2-associated syndrome presenting with Rett-like phenotypes. *Clin. Genet.* **2016**, *89*, 728–732. [CrossRef]
36. Saez, M.A.; Fernandez-Rodriguez, J.; Moutinho, C.; Sanchez-Mut, J.V.; Gomez, A.; Vidal, E.; Petazzi, P.; Szczesna, K.; Lopez-Serra, P.; Lucariello, M.; et al. Mutations in JMJD1C are involved in Rett syndrome and intellectual disability. *Genet. Med.* **2016**, *18*, 378–385. [CrossRef]
37. Rocha, H.; Sampaio, M.; Rocha, R.; Fernandes, S.; Leao, M. MEF2C haploinsufficiency syndrome: Report of a new MEF2C mutation and review. *Eur. J. Med. Genet.* **2016**, *59*, 478–482. [CrossRef]
38. Lopes, F.; Barbosa, M.; Ameur, A.; Soares, G.; de Sa, J.; Dias, A.I.; Oliveira, G.; Cabral, P.; Temudo, T.; Calado, E.; et al. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J. Med. Genet.* **2016**, *53*, 190–199. [CrossRef]
39. Sajan, S.A.; Jhangiani, S.N.; Muzny, D.M.; Gibbs, R.A.; Lupski, J.R.; Glaze, D.G.; Kaufmann, W.E.; Skinner, S.A.; Annese, F.; Friez, M.J.; et al. Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet. Med.* **2017**, *19*, 13–19. [CrossRef]
40. Allou, L.; Julia, S.; Amsallem, D.; El Chehadeh, S.; Lambert, L.; Thevenon, J.; Duffourd, Y.; Saunier, A.; Bouquet, P.; Pere, S.; et al. Rett-like phenotypes: Expanding the genetic heterogeneity to the KCNA2 gene and first familial case of CDKL5-related disease. *Clin. Genet.* **2017**, *91*, 431–440. [CrossRef]
41. Yoo, Y.; Jung, J.; Lee, Y.N.; Lee, Y.; Cho, H.; Na, E.; Hong, J.; Kim, E.; Lee, J.S.; Lee, J.S.; et al. GABBR2 mutations determine phenotype in rett syndrome and epileptic encephalopathy. *Ann. Neurol.* **2017**, *82*, 466–478. [CrossRef]

42. Vuillaume, M.L.; Jeanne, M.; Xue, L.; Blesson, S.; Denomme-Pichon, A.S.; Alirol, S.; Brulard, C.; Colin, E.; Isidor, B.; Gilbert-Dussardier, B.; et al. A novel mutation in the transmembrane 6 domain of GABBR2 leads to a Rett-like phenotype. *Ann. Neurol.* **2018**, *83*, 437–439. [CrossRef]
43. Huisman, S.; Mulder, P.A.; Redeker, E.; Bader, I.; Bisgaard, A.M.; Brooks, A.; Cereda, A.; Cinca, C.; Clark, D.; Cormier-Daire, V.; et al. Phenotypes and genotypes in individuals with SMC1A variants. *Am. J. Med. Genet. A* **2017**, *173*, 2108–2125. [CrossRef]
44. Percy, A.K.; Lane, J.; Annese, F.; Warren, H.; Skinner, S.A.; Neul, J.L. When Rett syndrome is due to genes other than MECP2. *Transl. Sci. Rare Dis.* **2018**, *3*, 49–53. [CrossRef]
45. Iwama, K.; Mizuguchi, T.; Takeshita, E.; Nakagawa, E.; Okazaki, T.; Nomura, Y.; Iijima, Y.; Kajiura, I.; Sugai, K.; Saito, T.; et al. Genetic landscape of Rett syndrome-like phenotypes revealed by whole exome sequencing. *J. Med. Genet.* **2019**. [CrossRef]
46. Vidal, S.; Brandi, N.; Pacheco, P.; Maynou, J.; Fernandez, G.; Xiol, C.; Pascual-Alonso, A.; Pineda, M.; Rett Working, G.; Armstrong, J. The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome. *Eur. J. Paediatr. Neurol.* **2019**. [CrossRef]
47. Swanson, D.A.; Steel, J.M.; Valle, D. Identification and characterization of the human ortholog of rat STXBP1, a protein implicated in vesicle trafficking and neurotransmitter release. *Genomics* **1998**, *48*, 373–376. [CrossRef]
48. Toonen, R.F.; Wierda, K.; Sons, M.S.; de Wit, H.; Cornelisse, L.N.; Brussaard, A.; Plomp, J.J.; Verhage, M. Munc18-1 expression levels control synapse recovery by regulating readily releasable pool size. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18332–18337. [CrossRef]
49. Medrihan, L.; Tantalaki, E.; Aramuni, G.; Sargsyan, V.; Dudanova, I.; Missler, M.; Zhang, W. Early defects of GABAergic synapses in the brain stem of a MeCP2 mouse model of Rett syndrome. *J. Neurophysiol.* **2008**, *99*, 112–121. [CrossRef]
50. Sepp, M.; Pruunsild, P.; Timmusk, T. Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* **2012**, *21*, 2873–2888. [CrossRef]
51. de Winter, C.F.; Baas, M.; Bijlsma, E.K.; van Heukelingen, J.; Routledge, S.; Hennekam, R.C. Phenotype and natural history in 101 individuals with Pitt-Hopkins syndrome through an internet questionnaire system. *Orphanet J. Rare Dis.* **2016**, *11*, 37. [CrossRef]
52. Sanders, S.J.; Campbell, A.J.; Cottrell, J.R.; Moller, R.S.; Wagner, F.F.; Auldridge, A.L.; Bernier, R.A.; Catterall, W.A.; Chung, W.K.; Empfield, J.R.; et al. Progress in Understanding and Treating SCN2A-Mediated Disorders. *Trends Neurosci.* **2018**, *41*, 442–456. [CrossRef]
53. Proikas-Cezanne, T.; Waddell, S.; Gaugel, A.; Frickey, T.; Lupas, A.; Nordheim, A. WIPI-1alpha (WIPI49), a member of the novel 7-bladed WIPI protein family, is aberrantly expressed in human cancer and is linked to starvation-induced autophagy. *Oncogene* **2004**, *23*, 9314–9325. [CrossRef]
54. Haack, T.B.; Hogarth, P.; Kruer, M.C.; Gregory, A.; Wieland, T.; Schwarzmayr, T.; Graf, E.; Sanford, L.; Meyer, E.; Kara, E.; et al. Exome sequencing reveals de novo WDR45 mutations causing a phenotypically distinct, X-linked dominant form of NBIA. *Am. J. Hum. Genet.* **2012**, *91*, 1144–1149. [CrossRef]
55. Ohba, C.; Nabatame, S.; Iijima, Y.; Nishiyama, K.; Tsurusaki, Y.; Nakashima, M.; Miyake, N.; Tanaka, F.; Ozono, K.; Saitsu, H.; et al. De novo WDR45 mutation in a patient showing clinically Rett syndrome with childhood iron deposition in brain. *J. Hum. Genet.* **2014**, *59*, 292–295. [CrossRef]
56. Zweier, M.; Gregor, A.; Zweier, C.; Engels, H.; Sticht, H.; Wohlleber, E.; Bijlsma, E.K.; Holder, S.E.; Zenker, M.; Rossier, E.; et al. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum. Mutat.* **2010**, *31*, 722–733. [CrossRef]
57. Bienvenu, T.; Diebold, B.; Chelly, J.; Isidor, B. Refining the phenotype associated with MEF2C point mutations. *Neurogenetics* **2013**, *14*, 71–75. [CrossRef]
58. Guy, J.; Hendrich, B.; Holmes, M.; Martin, J.E.; Bird, A. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.* **2001**, *27*, 322–326. [CrossRef]
59. Chapleau, C.A.; Calfa, G.D.; Lane, M.C.; Albertson, A.J.; Larimore, J.L.; Kudo, S.; Armstrong, D.L.; Percy, A.K.; Pozzo-Miller, L. Dendritic spine pathologies in hippocampal pyramidal neurons from Rett syndrome brain and after expression of Rett-associated MECP2 mutations. *Neurobiol. Dis.* **2009**, *35*, 219–233. [CrossRef]
60. Guy, J.; Gan, J.; Selfridge, J.; Cobb, S.; Bird, A. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **2007**, *315*, 1143–1147. [CrossRef]

61. Moretti, P.; Levenson, J.M.; Battaglia, F.; Atkinson, R.; Teague, R.; Antalffy, B.; Armstrong, D.; Arancio, O.; Sweatt, J.D.; Zoghbi, H.Y. Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. *J. Neurosci.* **2006**, *26*, 319–327. [CrossRef]
62. Samaco, R.C.; Mandel-Brehm, C.; Chao, H.T.; Ward, C.S.; Fyffe-Maricich, S.L.; Ren, J.; Hyland, K.; Thaller, C.; Maricich, S.M.; Humphreys, P.; et al. Loss of MeCP2 in aminergic neurons causes cell-autonomous defects in neurotransmitter synthesis and specific behavioral abnormalities. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 21966–21971. [CrossRef]
63. Harrington, A.J.; Raissi, A.; Rajkovich, K.; Berto, S.; Kumar, J.; Molinaro, G.; Raduazzo, J.; Guo, Y.; Loerwald, K.; Konopka, G.; et al. MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders. *Elife* **2016**, *5*. [CrossRef]
64. Berret, E.; Barron, T.; Xu, J.; Debner, E.; Kim, E.J.; Kim, J.H. Oligodendroglial excitability mediated by glutamatergic inputs and Nav1.2 activation. *Nat. Commun.* **2017**, *8*, 557. [CrossRef]
65. Shin, W.; Kweon, H.; Kang, R.; Kim, D.; Kim, K.; Kang, M.; Kim, S.Y.; Hwang, S.N.; Kim, J.Y.; Yang, E.; et al. Scn2a Haploinsufficiency in Mice Suppresses Hippocampal Neuronal Excitability, Excitatory Synaptic Drive, and Long-Term Potentiation, and Spatial Learning and Memory. *Front. Mol. Neurosci.* **2019**, *12*, 145. [CrossRef]
66. Santos, T.C.; Wierda, K.; Broeke, J.H.; Toonen, R.F.; Verhage, M. Early Golgi Abnormalities and Neurodegeneration upon Loss of Presynaptic Proteins Munc18-1, Syntaxin-1, or SNAP-25. *J. Neurosci.* **2017**, *37*, 4525–4539. [CrossRef]
67. Li, H.; Zhu, Y.; Morozov, Y.M.; Chen, X.; Page, S.C.; Rannals, M.D.; Maher, B.J.; Rakic, P. Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry* **2019**, *24*, 1235–1246. [CrossRef]
68. Wan, H.; Wang, Q.; Chen, X.; Zeng, Q.; Shao, Y.; Fang, H.; Liao, X.; Li, H.S.; Liu, M.G.; Xu, T.L.; et al. WDR45 contributes to neurodegeneration through regulation of ER homeostasis and neuronal death. *Autophagy* **2019**. [CrossRef]
69. Ehrhart, F.; Coort, S.L.; Cirillo, E.; Smeets, E.; Evelo, C.T.; Curfs, L.M. Rett syndrome - biological pathways leading from MECP2 to disorder phenotypes. *Orphanet J. Rare Dis.* **2016**, *11*, 158. [CrossRef]
70. Ehrhart, F.; Coort, S.L.; Eijssen, L.; Cirillo, E.; Smeets, E.E.; Bahram Sangani, N.; Evelo, C.T.; Curfs, L.M.G. Integrated analysis of human transcriptome data for Rett syndrome finds a network of involved genes. *World J. Biol. Psychiatry* **2019**. [CrossRef]
71. Lekman, A.Y.; Hagberg, B.A.; Svennerholm, L.T. Membrane cerebral lipids in Rett syndrome. *Pediatr. Neurol.* **1991**, *7*, 186–190. [CrossRef]
72. Kishi, N.; MacDonald, J.L.; Ye, J.; Molyneaux, B.J.; Azim, E.; Macklis, J.D. Reduction of aberrant NF-kappaB signalling ameliorates Rett syndrome phenotypes in Mecp2-null mice. *Nat. Commun.* **2016**, *7*, 10520. [CrossRef]
73. di Michele, F.; Luchetti, S.; Bernardi, G.; Romeo, E.; Longone, P. Neurosteroid and neurotransmitter alterations in Parkinson's disease. *Front. Neuroendocrinol.* **2013**, *34*, 132–142. [CrossRef]
74. Goffin, D.; Brodkin, E.S.; Blendy, J.A.; Siegel, S.J.; Zhou, Z. Cellular origins of auditory event-related potential deficits in Rett syndrome. *Nat. Neurosci.* **2014**, *17*, 804–806. [CrossRef]
75. Chao, H.T.; Chen, H.; Samaco, R.C.; Xue, M.; Chahrour, M.; Yoo, J.; Neul, J.L.; Gong, S.; Lu, H.C.; Heintz, N.; et al. Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature* **2010**, *468*, 263–269. [CrossRef]
76. Banerjee, A.; Rikhye, R.V.; Breton-Provencher, V.; Tang, X.; Li, C.; Li, K.; Runyan, C.A.; Fu, Z.; Jaenisch, R.; Sur, M. Jointly reduced inhibition and excitation underlies circuit-wide changes in cortical processing in Rett syndrome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E7287–E7296. [CrossRef]
77. Tang, X.; Kim, J.; Zhou, L.; Wengert, E.; Zhang, L.; Wu, Z.; Carromeu, C.; Muotri, A.R.; Marchetto, M.C.; Gage, F.H.; et al. KCC2 rescues functional deficits in human neurons derived from patients with Rett syndrome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 751–756. [CrossRef]
78. Cosentino, L.; Vigli, D.; Franchi, E; Laviola, G.; De Filippis, B. Rett syndrome before regression: A time window of overlooked opportunities for diagnosis and intervention. *Neurosci. Biobehav. Rev.* **2019**. [CrossRef]
79. Bedogni, F.; Cobolli Gigli, C.; Pozzi, D.; Rossi, R.L.; Scaramuzza, L.; Rossetti, G.; Pagani, M.; Kilstrup-Nielsen, C.; Matteoli, M.; Landsberger, N. Defects During Mecp2 Null Embryonic Cortex Development Precede the Onset of Overt Neurological Symptoms. *Cereb Cortex* **2016**, *26*, 2517–2529. [CrossRef]

80. Smith-Hicks, C.L.; Gupta, S.; Ewen, J.B.; Hong, M.; Kratz, L.; Kelley, R.; Tierney, E.; Vaurio, R.; Bibat, G.; Sanyal, A.; et al. Randomized open-label trial of dextromethorphan in Rett syndrome. *Neurology* **2017**, *89*, 1684–1690. [CrossRef]
81. Mancini, J.; Dubus, J.C.; Jouve, E.; Roux, J.C.; Franco, P.; Lagrue, E.; Castelnau, P.; Cances, C.; Chaix, Y.; Rougeot-Jung, C.; et al. Effect of desipramine on patients with breathing disorders in RETT syndrome. *Ann. Clin. Transl. Neurol.* **2018**, *5*, 118–127. [CrossRef]
82. O'Leary, H.M.; Kaufmann, W.E.; Barnes, K.V.; Rakesh, K.; Kapur, K.; Tarquinio, D.C.; Cantwell, N.G.; Roche, K.J.; Rose, S.A.; Walco, A.C.; et al. Placebo-controlled crossover assessment of mecasermin for the treatment of Rett syndrome. *Ann. Clin. Transl. Neurol.* **2018**, *5*, 323–332. [CrossRef]
83. Stamberger, H.; Weckhuysen, S.; De Jonghe, P. STXBP1 as a therapeutic target for epileptic encephalopathy. *Expert. Opin. Ther. Targets* **2017**, *21*, 1027–1036. [CrossRef]
84. Gerber, S.H.; Rah, J.C.; Min, S.W.; Liu, X.; de Wit, H.; Dulubova, I.; Meyer, A.C.; Rizo, J.; Arancillo, M.; Hammer, R.E.; et al. Conformational switch of syntaxin-1 controls synaptic vesicle fusion. *Science* **2008**, *321*, 1507–1510. [CrossRef]
85. Monteggia, L.M.; Kavalali, E.T. Rett syndrome and the impact of MeCP2 associated transcriptional mechanisms on neurotransmission. *Biol. Psychiatry* **2009**, *65*, 204–210. [CrossRef]
86. Dani, V.S.; Nelson, S.B. Intact long-term potentiation but reduced connectivity between neocortical layer 5 pyramidal neurons in a mouse model of Rett syndrome. *J. Neurosci.* **2009**, *29*, 11263–11270. [CrossRef]
87. Chamma, I.; Chevy, Q.; Poncer, J.C.; Levi, S. Role of the neuronal K-Cl co-transporter KCC2 in inhibitory and excitatory neurotransmission. *Front. Cell Neurosci.* **2012**, *6*, 5. [CrossRef]
88. Duarte, S.T.; Armstrong, J.; Roche, A.; Ortez, C.; Perez, A.; O'Callaghan Mdel, M.; Pereira, A.; Sanmarti, F.; Ormazabal, A.; Artuch, R.; et al. Abnormal expression of cerebrospinal fluid cation chloride cotransporters in patients with Rett syndrome. *PLoS ONE* **2013**, *8*, e68851. [CrossRef]
89. Lozovaya, N.; Nardou, R.; Tyzio, R.; Chiesa, M.; Pons-Bennaceur, A.; Eftekhari, S.; Bui, T.T.; Billon-Grand, M.; Raserom, J.; Bonifazi, P.; et al. Early alterations in a mouse model of Rett syndrome: The GABA developmental shift is abolished at birth. *Sci. Rep.* **2019**, *9*, 9276. [CrossRef]
90. Tang, X.; Drotar, J.; Li, K.; Clairmont, C.D.; Brumm, A.S.; Sullins, A.J.; Wu, H.; Liu, X.S.; Wang, J.; Gray, N.S.; et al. Pharmacological enhancement of KCC2 gene expression exerts therapeutic effects on human Rett syndrome neurons and Mecp2 mutant mice. *Sci. Transl. Med.* **2019**, *11*, eaau0164. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Capítulo 4

Capítulo 4: Caracterización funcional de hallazgos detectados por NGS.

En la búsqueda de nuevos genes que puedan estar asociados al fenotipo RTT/RTT-*like*, no solamente encontramos genes candidatos que ya han sido descritos para otros trastornos del neurodesarrollo en los cuales ya hay un extenso conocimiento sobre las funciones que desempeñan y que problemas acarrean su mal funcionamiento en el organismo. Pero también detectamos mutaciones con una alta probabilidad de que sean patogénicas en genes que no han sido asociados a ninguna enfermedad y que apenas se conocen las funciones que desarrollan en el organismo. De ese modo, su caracterización funcional es estrictamente necesaria para poder relacionarlos con certeza al fenotipo de las pacientes. En este capítulo presentamos los estudios funcionales realizados para una de las variantes candidatas detectadas en los WES realizados en pacientes con RTT/RTT-*like* en el primer trabajo.

Inicialmente fueron elegidas e iniciados estudios funcionales para dos variantes candidatas en los genes *SLC6A1* y *GABBR2*, ambos implicados en la vía GABAérgica (Fig 10). Fueron elegidos por su implicación en esta vía, que ya se conocía su alteración en el RTT desde antes que se conociera el gen *MECP2* como causante de la enfermedad. El gen *SLC6A1* codifica para uno de los principales transportadores de GABA del SNC. GAT-1 se expresa en los terminales nerviosos de las neuronas GABAérgicas, pero también se puede encontrar en los astrocitos, y se ha demostrado que es crucial para la recaptación de GABA de la sinapsis. Por otro lado, el gen *GABBR2* codifica para la subunidad 2 de un receptor metabotrópico que inhibe la actividad neuronal a través de los sistemas de segundo mensajero acoplados a la proteína G, que regulan la liberación de neurotransmisores y la actividad de los canales iónicos y la adenil ciclasa (Dore et al., 2014). Esta subunidad del receptor forma un complejo heterodimérico activo con la subunidad del receptor GABAB (1), ninguno de los cuales es eficaz por sí solo. Ambos son necesarios para el correcto funcionamiento de la proteína de membrana. En el receptor heteromérico, GABAB (1) es responsable de la unión de GABA, mientras que GABAB (2) es necesario para el tráfico de superficie y el acoplamiento de la proteína G. La activación de GABABR induce la disociación de Gi / o trimérica unida al receptor en las subunidades Gai / o y Gbc, y la liberación de Gai / o inhibe la adenilil ciclasa, disminuyendo el nivel de AMP cíclico intracelular (cAMP), un regulador de la proteína quinasa A. A través de esto acción,

GABABR ejerce efectos lentos y duraderos sobre el desarrollo neural, manteniendo un delicado equilibrio de señalización neuronal excitadora e inhibitoria (Wu et al., 2014).

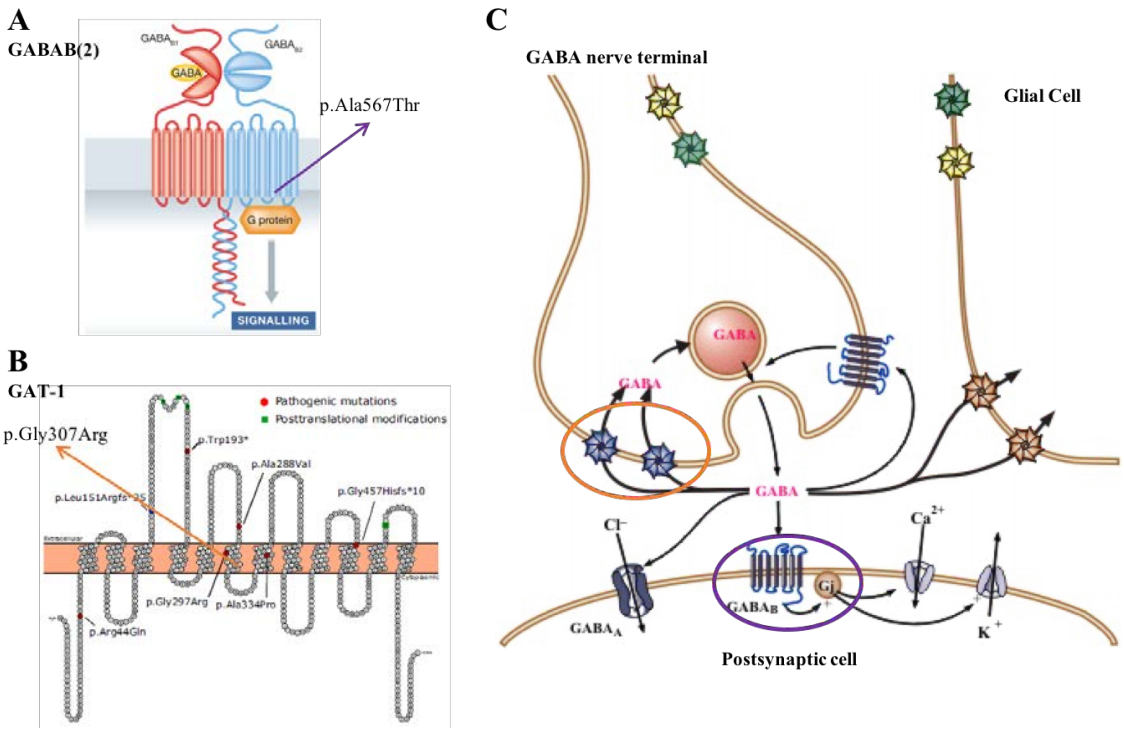


Figura 10. Representación y localización de las proteínas codificadas por los genes *GABBR2* y *SLC6A1*. A) Proteína GABAB(2), en azul, codificada por *GABBR2*. B) Proteína GAT-1 codificada por *SLC6A1*. C) Esquema de la localización y circulación del GABA en la sinapsis, en morado se encuentra marcado el receptor GABAB(2) y en naranja el transportador GAT-1. *Imágenes modificadas de Benke, D. et al (2012), Carvill, G.L. et al (2015) y Arbeitskereis Prof. Dr. Klaus T. Wanner Lehrstuhl für Pharmazeutische Chemie. Ludwig-Maximilians-Universität München, respectivamente.*

Los estudios se iniciaron con análisis *in silico* de las proteínas mutantes seguido de la preparación de los vectores tanto salvaje como mutante para realizar ensayos *in vitro* en líneas celulares. Sin embargo, durante el transcurso de los experimentos se publicó un extenso estudio acerca de la patogenicidad del cambio que detectamos en *GABBR2* confirmando la patogenicidad del cambio y su relación con el fenotipo RTT-*like* que presentaba su paciente (Yoo et al., 2017). En este punto, se decidió para los estudios que se querían realizar en nuestro laboratorio con el gen *GABBR2* ya que se confirma claramente que la variante es patogénica y causante de la enfermedad de nuestra paciente también. No obstante, los estudios funcionales para el gen *SLC6A1* continuaron su transcurso y se encuentran resumidos en nuestra última publicación.

Publicación 6

Título: *A de-novo SLC6A1 gene mutation in a patient showing the Rett-like syndrome phenotype***.**

Autores: Silvia Vidal, Mónica Roldan, Mercedes Casado, Janet Hoenicka, Mercè Pineda, Judith Armstrong.

Referencia: *Clinical Genetics* **(**En preparación)

Resumen:

En este trabajo presentamos el primer caso de un paciente RTT-*like* en España que presenta la variante (p.G307R) en el gen *SLC6A1* y los primeros estudios funcionales sobre su patogenicidad. Este gen se describió por primera vez en 2015 asociado con epilepsia atónica mioclónica y, más adelante, con discapacidad intelectual con epilepsia y, a veces, con regresión del lenguaje. Estas descripciones clínicas previas y la certeza que tenemos que en las pacientes con RTT la vía GABAérgica está alterada, procedemos al análisis tanto *in silico* como *in vitro* de la variante detectada. Realizamos, en un primer paso, estudios *in silico* que nos permiten predecir las posibles implicaciones a nivel estructural que pueda estar causando la mutación en el correcto plegamiento y funcionalidad de la proteína. Además de, realizar estudios *in vitro* en líneas celulares estables, para identificar si la proteína mutante es capaz de llegar a la membrana citoplasmática y es capaz de realizar su función.

A de-novo *SLC6A1* gene mutation in a patient showing the Rett-like syndrome phenotype

Silvia Vidal[a,b], Mónica Roldan[c], Mercedes Casado[d], Janet Hoenicka[b,e], Mercè Pineda[a], Judith Armstrong[f].

[a]Sant Joan de Déu Research Foundation, 08950 Barcelona, Spain.
[b]IPER, Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, 08950 Barcelona, Spain.
[c]Confocal Microscopy Unit, Department of Pathology, Hospital Sant Joan de Déu, Barcelona, Spain.
[d]Clinical Biochemistry Department, Hospital Sant Joan de Déu, Barcelona, Spain; CIBERER-Instituto de Salud Carlos III, Barcelona, Spain.
[e]Laboratory of Neurogenetics and Molecular Medicine; Centro de Investigación Biomédica en Red de Salud Mental (CIBESAM), Barcelona, Spain.
[f]Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain; Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain; CIBER-ER (Biomedical Network Research Center for Rare Diseases), Institute of Health Carlos III (ISCIII), Madrid, Spain.

Correspondence:

Judith Armstrong
Molecular and Genetics Medicine Section, Hospital Sant Joan de Déu, Barcelona, Spain; Institut de Recerca Pediàtrica Hospital Sant Joan de Déu, Barcelona, Spain; CIBER-ER (Biomedical Network Research Center for Rare Diseases), Institute of Health Carlos III (ISCIII), Madrid, Spain. Electronic address: jarmstrong@sjdhospitalbarcelona.org.

ACKNOWLEDGEMENTS

We thank all patients and their families who contributed to this study. The work was supported by grants from the Spanish Ministry of Health (Instituto de Salud Carlos III/FEDER, PI15/01159); Crowdfunding program PRECIPITA, from the Spanish Ministry of Health (Fundación Española para la Ciencia y la Tecnología); the Catalan Association for Rett Syndrome; Fondobiorett and Mi Princesa Rett.

CONFLICT OF INTEREST STATEMENT

None of the authors have any conflict of interest to declare.

DATA AVAILABILITY STATEMENT

All data from this article is available in the Supplementary Data.

Abstract

Rett syndrome (RTT) is an early-onset neurodevelopmental disorder that primarily affects females, resulting in severe cognitive and physical disabilities, and is one of the most prevalent causes of intellectual disability in females. This neurodevelopmental disorder mostly caused by mutations in Methyl-CpG-binding protein *2* (*MECP2*); however, mutations in various other genes may lead to RTT-like phenotypes. The Solute Carrier Family 6 Member 1 (*SLC6A1*) gene was first reported in 2015 associated with myoclonic atonic epilepsy (MAE) and intellectual disability (ID) with epilepsy and sometimes with language regression. Here, we report the first case of RTT-like patient in Spain that presents the variant (p.G307R) in *SLC6A1* gene and the firsts preliminary functional studies about its pathogenicity.

Keywords: Rett syndrome, GAT-1, GABA transporter, functional study

1. Introduction

Rett syndrome (RTT, OMIM#312750) is an early-onset neurodevelopmental disorder characterized by a period of normal development until the first 6-18 months of life followed by a regression of acquired psychomotor abilities. RTT features include severe mental retardation, epilepsy, regression of purposeful hand use and language, breathing disturbances, gait apraxia and repetitive stereotyped hand movements (Neul et al., 2010; Weaving, Ellaway, Gecz, & Christodoulou, 2005). RTT has an incidence of 1:10,000–20,000 live births and affects mainly young females, being the second most common cause of severe mental retardation in females after Down syndrome (Liyanage & Rastegar, 2014).

Although the majority of RTT patients have pathogenic variants in the gene encoding methyl-CpG binding protein 2 (MECP2, OMIM*300005), approximately 5% of classic RTT and 25% of atypical RTT patients are negative for *MECP2* pathogenic variants However, the aetiology of some of RTT patients still remains unknown. (Neul et al., 2010; Neul et al., 2014). MeCP2 is a transcriptional regulatory protein, and in its absence, a large number of genes exhibit abnormal expression with implications in the balance between synaptic excitation and inhibition (Chao et al., 2010).

In recent years, the next generation sequencing (NGS) studies performed with patients within the RTT spectrum have revealed a wide range of candidate genes to explain the clinical nature of this patient cohort (Vidal et al., 2019). One of these new candidate gene is the solute carrier family member 1 gene (*SLC6A1*; OMIM*137165), related with the GABAergic pathway. This pathway is altered in RTT patients, for example, deletion of Mecp2 in γ-aminobutyric acid (GABA)-expressing neurons results in reduced GABA signalling and recapitulates most of the RTT features in mice (Chao et al., 2010), whereas the ablation of Mecp2 in excitatory glutamatergic neurons caused neurological features different from those of RTT (Meng et al., 2016). The *SLC6A1* gene encodes the gamma-aminobutyric acid (GABA) transporter 1 (GAT-1), which removes GABA from the synaptic cleft restoring it to presynaptic terminals and is one of the main GABA transporters (Zafar & Jabeen, 2018). The clinical picture of *SLC6A1* gene mutations is characterized by a broader spectrum including a mild-to-moderate intellectual disability, speech difficulties, behavioural problems (such as hyperactivity, attention deficit, aggressiveness, and autistic traits), epilepsy (often with myoclonic-atonic and atypical absence seizures, characterizing a myoclonic-atonic epilepsy), and neurological signs (ataxia or unsteady gait, hypotonia, tremor, and fine-motor impairment) (Carvill et al., 2015; Johannesen et al., 2018). Given the previous observation that *SLC6A1* is also mutated in patients with a similar phenotype as RTT-like patients, we explored the molecular consequences of our *SLC6A1* mutation using *in silico* analysis and cell culture and found that the severity of the variant could determine the phenotypic outputs.

2. Material and Methods

2.1 Molecular analysis and in silico analysis

Written informed consent was obtained for molecular analyses. The whole exome sequencing by trios was carried out has been described in the article Lucariello et al. 2016. We use different *in silico* prediction softwares: Polyphen2 (http://genetics.bwh.harvard.edu/pph2/), SIFT (https://sift.bii.a-star.edu.sg/) and Mutation Taster (http://www.mutationtaster.org/), to predict variants as damaging or harmless. Furthermore, we took into account the results observed in ClinVar, that aggregates information about genomic variation and its relationship to human health, and Varsome, variant knowledge community, data aggregator and variant data discovery tool (Kopanos et al., 2019).

Finally, it has been use software Phyre2, web portal for protein modelling, prediction and analysis, to obtain the protein structure prediction (Kelley, Mezulis, Yates, Wass, & Sternberg, 2015). And the 93% of residues modelled at >90% confidence. And the Missense 3D, tool that software that predicts the structural changes introduced by an amino acid substitution and is applicable to analyse both PDB coordinates and homology-predicted structures (Ittisoponpisan et al., 2019).

2.2 in vitro assay

For the *in vitro* studies in cell lines we designed and cloned, in the green fluorescence protein (GFP)-tagged expression vector, the wild-type (WT) gene and the patient's variant was obtained by site-directed mutagenesis (QuickChange Lightning Site-Directed kit, Aligent Technologies). These constructs were transfected with FuGene (Promega) in human embryonic kidney (HEK293T) cells. Cell counting for transfection rate was performed with Leica DMI 3000B microscope (Leica Microsystems) and 63x oil immersion objective and five independent assays were

counted. We studied subcellular localization and expression pattern by immunofluorescence staining assay.

2.2.1 Immunofluorescence Assay

HEK293T cells transfected with SLC6A1 mutated (G307R) and wildtype (WT) constructs were cultured onto poly-L-lysine (Sigma-Aldrich, St. Louis, MO) coated coverslips, fixed with 4% paraformaldehyde for 20 minutes at 37C, and then washed briefly in PBS. Then, for cell membrane staining, samples were incubated with WGA for 10 minutes, and then washed briefly in PBS (Wheat Germ Agglutinin, Alexa Fluor™ 594 Conjugate, W11262, Invitrogen). Cells were permeabilized with 0.1% Saponin PBS for 10 minutes and then were washed with PBS three times. Cells were blocked with 1% bovine serum albumin in PBS for 60 minutes and then incubated in primary antibody anti-GM130 for Golgi staining and anti-Calreticulina for endoplasmatic reticulum staining (anti-GM130, 1:200 (ab169276, Abcam); anti-Careticulina, 1:150 (PA3-900; Thermo Fisher]) overnight at 4ºC. Samples were washed with PBS-Tween three times, then incubated with secondary antibodies (Goat Anti-Mouse IgG H&L (Texas Red ®) 1:2000 (ab6787, Abcam); Donkey Anti-Mouse IgG H&L (Alexa Fluor® 594), 1:2000 (ab150108; Abcam), for 45 minutes at room temperature. Then, mounted on slides using Fluoromount-G™ Mounting Medium, with DAPI (00-4959-52, Invitrogen).

Cell counting was performed with Leica DMI 3000B microscope (Leica Microsystems) and 63x oil immersion objective. Five independent assays were counted for each vector (WT and G307R), at least 100 random cells per assay. High resolution images were acquired with a Leica TCS SP8 X confocal microscope (Leica Microsystems).

2.2.2 Cycloheximide Chase Assay

It has been studied the protein stability of the protein WT and G307R in order to determine the half-life of a given protein. HEK293T were cultured 75×10^4 cells in 35-mm dishes and have them incubated in a CO_2 incubator overnight and transfected with SLC6A1 G307R and WT constructs. 72h post-transfected cells where incubated with Complete Media with Cycloheximide (50ug/ml) during 2, 4, 6 and 8 hours. Then, cells where collected and the protein were extracted. The protein extraction was analysed by western blotting and bands quantification by ImageJ.

3. Results

3.1 Clinical presentation

Our patient, 36-and-a-half-year-old woman, controlled in paediatric neurology department due to global delay with favourable evolution. The first signs of global delay appeared around 6 months of age, being assessed in paediatric neuropathy, and the clinical signs were initially interpreted as ataxic-spastic cerebral palsy. Progressively, hypotonia and lack of balance predominated, with frequent block falls. She acquired free running at 2 years, at which time seizures of absence and difficulties in sleep appeared, achieving control of the seizures with ethosuximide and valproate; subsequently ethosuximide could be removed. She started the language with loose words towards the age of 2. They refer to a functional and purposeful use of the hands, without the appearance of stereotypes. Given the persistence of an unaffiliated global delay, in a girl who presented a deterioration in the rate of development between 6 and 10 months, given the possibility of an atypical form of RTT. A study of *MECP2* and *CDKL5* was carried out in 2009, and cranial MRI were normal. She went to a special school, where she was sociable until 14-15 years old, at which time she presented a sudden change in character, with a tendency to self-harm, which improved with Zyprexa. At the age of 32, diabetes was diagnosed, currently being treated with metformin. Hypothyroidism was also identified, and treated with Eutirox.

3.2 Molecular genetics

No pathogenic variants were found in *MECP2*. Whole exome sequencing in the patient identified a *SLC6A1* mutation (NM_003042.3: c.919G>A), which induced an amino acid change (NP_003033.3: p.Gly307Arg). The c.919G>A variant is absented in the Genome Aggregation Database (gnomAD; gnomad.broadinstitute.org). This variant is classified in ClinVar as a likely pathogenic variant, because all of these evidences but it was considered that the possibility it may be a rare benign variant cannot be excluded. Based on the pathogenicity classification of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) guidelines, this variant is predicted to be likely pathogenic (PP2, PP3, PP5, and PM2). The WES results in trio and the Sanger sequencing confirmation of parental samples indicated that c.919G>A occurred *de novo* in the patient.

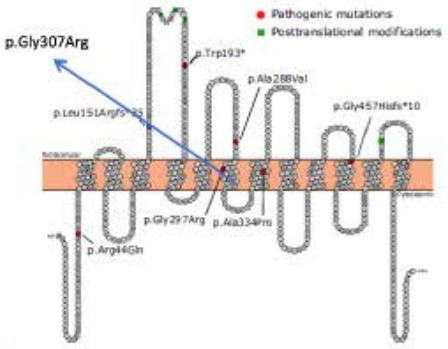
3.3 *in silico* analysis

In silico analysis predicts this variant is probably damaging to the protein structure/function. This substitution occurs at a position that is conserved across species and it is localized in a transmembrane domain of the protein (Fig. 1A). The Gly307 is a non-conservative amino acid substitution, which is likely to impact secondary protein structure as these residues differ in charge and size (Fig. 1B). In order to further study the effect that this amino acid change may be causing, we generated the GAT-1 protein structure prediction using Phyre2 (Fig 1C). Then, it has been introduced de amino acid change analysing it in the Missense 3D. Its results suggest that this change generate a structural damage, because of the change of polarity of amino acid and the fact that buried glycines are normally highly conserved. Moreover, it has been predicted a cavity alteration, the substitution leads to the contraction of cavity volume by 111.456 Å^3.

3.4 *in vitro* analysis

The transfection rate was 60% of total cells (data not shown). Through the results obtained in the immunofluorescence assays, it has been possible to detect that both the GAT-1 G307R and the WT protein can reach the cytoplasmic membrane, verifying this by means of a colocalization analysis (Fig. 2). However, there are tangible differences between them, the most significant being that in WT transfected cells there is a greater migration of the protein of interest to the membrane, based on both the number of cells and the quantity. Based on the distribution of the construct in the cell, the presence of two patterns of micrographs in WT transfected cells can be verified, membrane-reaching protein and cytosolic protein aggregates (Fig 3); and four in the case of G307R, membrane protein, cytosolic protein, non-aggregated protein that does not reach the membrane and protein forming a mesh in the cytoplasm (Fig 4). A manual count of the different cell morphologies was performed, obtaining that 60% (69

A



B

Protein change	SIFT Prediction		Polyphen Prediction		PROVEAN Prediction		Mutation taster Prediction	
	Score	Prediction (cutoff = 0.05)	Score	Prediction (cutoff = 0.9)	Score	Prediction (cutoff = -2.5)	Score	Prediction (cutoff = 0.98)
Missense	0.000	Deleterious	1	Damaging	-7.609	Deleterious	0.999	Disease causing

Species	match	aa	alignment
H. sapiens WT		307	GLGLGSLIALGSYNSFHNNVYRD
H. sapiens G307R	not conserved	307	GSLIALRSYNSFHNNVYR
M. mulatta	all identical	307	GSLIALGSYNSFHNNVYR
M. musculus	all identical	307	GSLIALGSYNSFHNNVYR
G. gallus	all identical	306	GSLIALGSYNPFHNNVYR
T. rubripes	all identical	310	LGLGSLIALGSYNTFNNNVYR
D. rerio	all identical	306	GSLIALGSYNPFNNNVYR
D. melanogaster	all identical	345	GLGLGTLVALGSYNKFTNNVYK
Celegans	all identical	316	GLGLGALVALGSYNKFNNNVYK
X. tropicalis	all identical	228	GLGLGSLIALGSYNRFHNNVYR

C

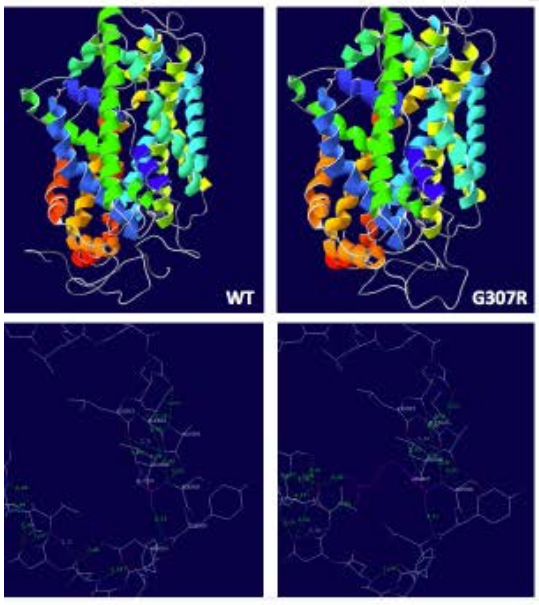


Figure 1. GAT-1 protein coded by *SLC6A1* gene and the variant analyzed in this study (p.G307R).

A) Schematic representation of GAT-1 transmembrane protein and the variant localized in this study localized with other pathogenic variants described; *modified from* Carvill, G.L. et al (2015).

B) Data results of prediction tools and the protein region conservation among species.

C) Image of GAT-1 protein structure prediction fold using Phyre2 software. In the left site WT protein and in the right site mutant protein.

of 115 observed cells) reach the membrane in WT cells and the 41% (49 of 120 observed cells) in G307R (Fig 5).

In order to study the stability of the WT and G307R protein, we carried out a cyclohexamide assay at different periods of times to determine if the half-life of both proteins differed. Figure 6 shows the results of said assay, and the development of the WB suggests further degradation of the mutant protein compared to WT.

3. Discussion

Variants in *SLC6A1* were first identified in patients presenting with myoclonic-atonic epilepsy, which is characterized by a range of seizure types including myoclonic, myoclonic-atonic, atonic, and absence seizures (Carvill et al., 2015; Mattison et al., 2018). Additionally, patients with myoclonic-atonic epilepsy also have variable degrees of intellectual disability and developmental delay, and in some cases, it has been also observed behavioural problems, namely, attention

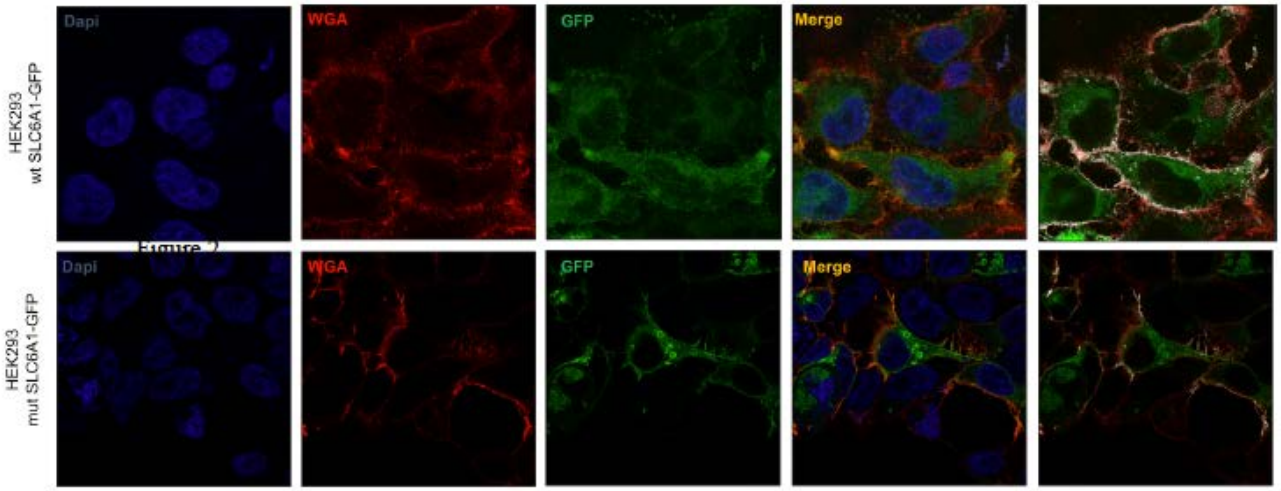


Figure 2. Confocal images of HEK293 transfected with GAT-1 WT and G07R mutant. Merge images shows the co-localization of the membrane with GAT-1 protein, both, WT and mutant, can reach the membrane. In the last two images the pixels that co-localise membrane with GAT-1 are coloured in white. Dapi-Nucleus staining; WGA-Membrane staining; GFP-GAT-1 protein.

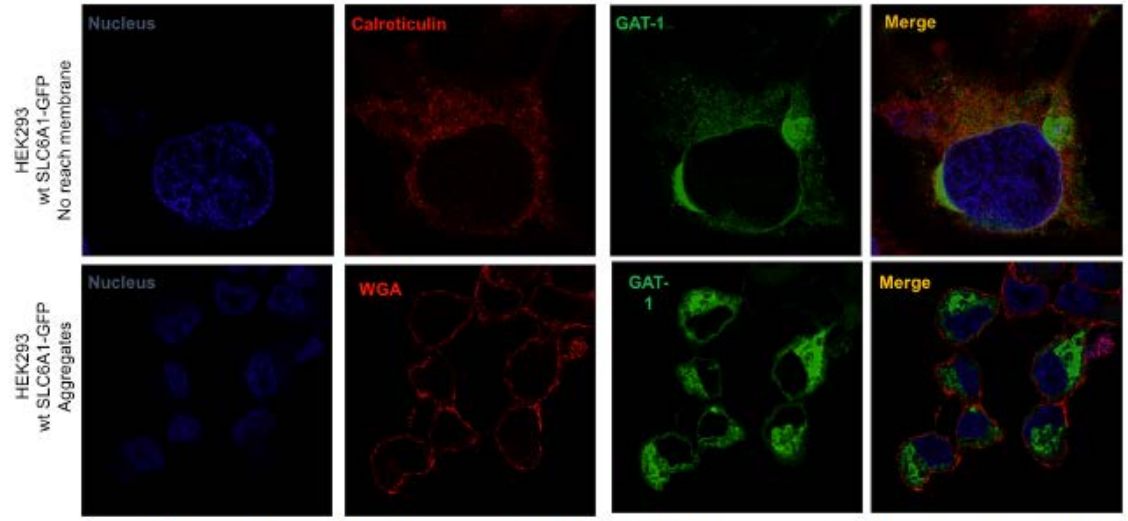


Figure 3. Confocal images of HEK293 transfected with GAT-1 WT. Merge images shows the co-localization of the ER with GAT-1 protein in cells that protein cannot reach membrane; and the aggregates of GFP inside the cells with apoptotic morfology; Dapi-Nuecleus staining; WGA-Membrane staining; GFP-GAT-1 protein.

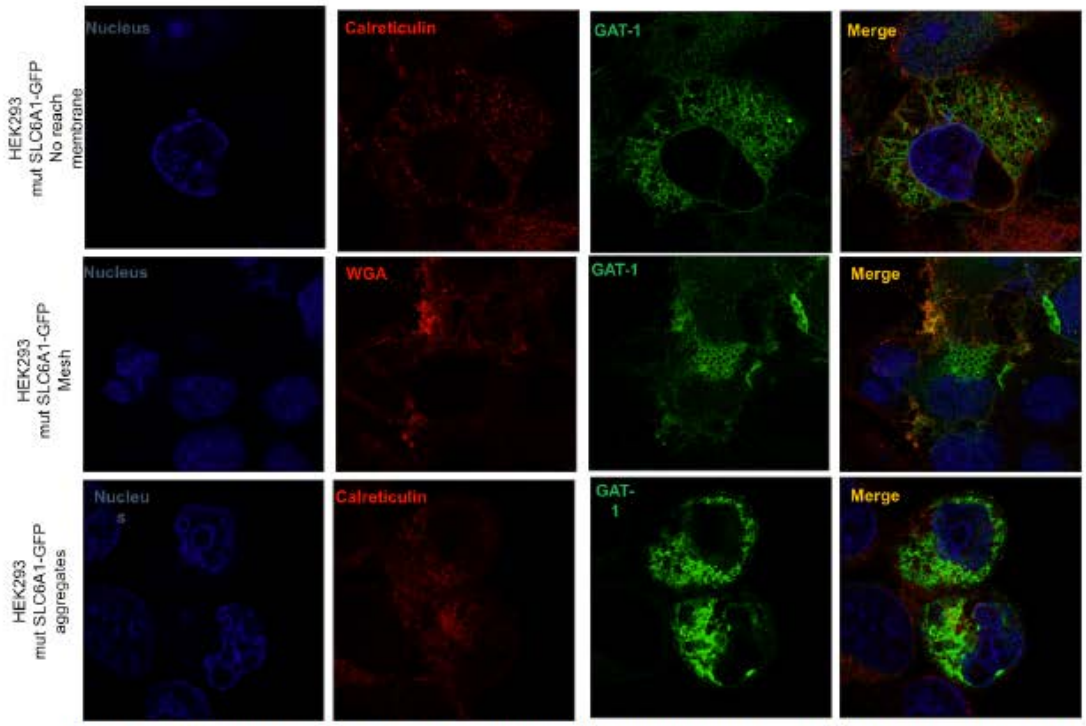


Figure 4. Confocal images of HEK293 transfected with GAT-1 G307R. Merge images shows the co-localization of the ER with GAT-1 protein in cells that GAT-1 no reach de membrane; the GAT-1 mesh that is form in some cells with normal appearance and the aggregates of GFP inside the cells with apoptotic form; Dapi-nucleus; WGA-Membrane staining; Calerticulin-ER staining; GFP-GAT-1 protein.

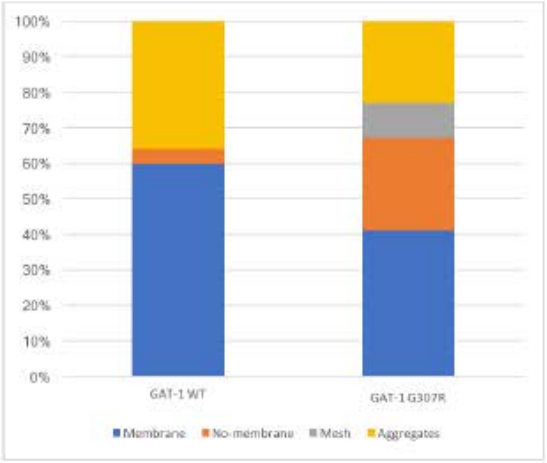


Figure 5. Graphic of manual count of the different cell morphologies detected inHEK293 cells transfected with GAT-1 WT and G07R.

deficit, hyperactivity, aggressive behaviour, hand stereotypies, and autistic features in different combinations (Johannesen et al., 2018). Furthermore, it has been described patients with autism spectrum disorder with language regression similar to our patient's phenotype (Islam, Herman, & de Los Reyes, 2018). Finally, focusing on the treatment to improve epilepsy in our patient, we have been able to see that the administration of valproate has given a good response. These results of the treatment with valproate is similar to those obtained in patients with a *SLC6A1* mutation were the epilepsy was controlled in the same way as in our patient (Johannesen et al., 2018; Posar & Visconti, 2019). This, in light of the mutation of the *SLC6A1* gene, was predictable as the valproic acid enhances the action of GABA by inhibiting its degradation and increasing its production (Chateauvieux, Morceau, Dicato, & Diederich, 2010).

At the molecular level, we can appreciate that the change is in a highly conserved glycine between species. Glycine is a very unique amino acid in that in contains a hydrogen as its side chain rather than a carbon, as is the case in all other amino acids. The ability of glycine to be more flexible at conformational level can reach parts of protein structures that are forbidden to all other amino acids (e.g. tight turns in structures) (Betts & Russell, 2003). In our case, glycine is replaced by an arginine, amino acid with electrically positive charged side chain. In this way, the Missense 3D program predicts that this change could generate a structural damage and a reduction of the protein cavity due to the long chain of arginine. Changes in cavity

volume can affect the protein stability or in this case could be affecting the hole of the transporter.

Our functional studies on the pathogenicity of this change in the *SLC6A1* gene provide us with a first approximation of what may be happening to the mutated protein in the cell. Cellular studies allow us to verify how both proteins (GAT-1 G307R and WT) reach the membrane, although it is feasible to infer that both the quantity and the quality of these differ, finding less protein in the mutant cells, at the same time as a g reater amount of intracellular protein aggregates. Aggregates that we also observe in WT cells, which may have their origin in a process of apoptosis due to

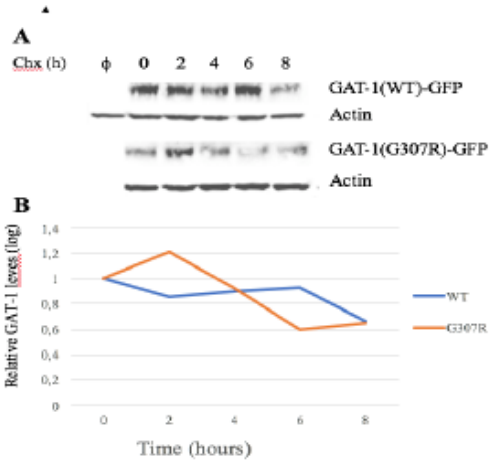


Figure 6. Effect of the mutation in GAT-1 protein stability was examined with cycloheximide-chase assay. A) Western blot analysis showed that treatment with cycloheximide affects the protein half-life. B) Graphic representation of protein level normalized with actin.

the morphological characteristics of these cells, rounded shape with no apparent intracellular organization. On the contrary, we observed how only in the case of GAT-1 G307R cells does the morphology appear in the form of a mesh, maintaining a healthy morphology. Following this reasoning and after correcting the cell counts by removing those that could be following an apoptotic process, we observed how in GAT-1 WT cells a high percentage of protein reaches the membrane, 94%; while in the case of GAT-1 G307R cells it reaches 61%.

During the study of the images, we were able to verify the existence of fluorescence aggregates in the cytosolic region of the mutant cells forming a mesh and more cells than WT type that cannot reach the membrane. There are several hypotheses that we can pose about the origin of this accumulation, being the one that convinced us the most an increase in the instability of the mutant protein. Performing the stability assay with cyclohexamide we are able to appreciate that GAT-1 G307R protein tends to degrade faster than WT. Thus, these fluorescence aggregates could constitute not only pools of mutant protein, but also pools of free or pseudo-free GFP due to digestion of the mutant protein.

More studies are needed to study the functionality of the channel. In an attempt to measure whether the GAT-1 channel is working properly, we performed the GABA uptake assay (Supplementary Data). The results obtained in this trial have not been conclusive, since we have not been able to appreciate differences between the functionality of the mutated channel with the WT (data not shown). It is probable that the GFP added to the N-terminal of the protein may be affecting the correct performance of the transporter and this leads to not observing differences in the functionality assay carried out on the transporter. New approaches have to be carried out to study the functionality of the channel. Removing the GFP label found in the design of our construct to have the protein in the same state as it would be in vivo, and creating stable cell lines that express it, could show us if there really are differences between the GAT-1 G307R protein and the GAT-1 WT.

In conclusion, *in silico* studies suggest that this change could be the cause of the disease in the patient. The protein is able to reach the membrane in its mutated form, although we cannot rule out a malfunction of the canal due to the decrease in the cavity predicted by the simulation. Immunofluorescent assays have been demonstrated that exist differences between both proteins, observing a lower arrival of mutant protein to the membrane, forming aggregates in the cytosol and with a lower stability compared to the WT protein.
These first in vitro functional studies suggest that there are differences between the functionality of the WT and the mutant protein, although new approaches have to be made to finally define the pathogenicity of the change.

REFERENCES

Betts, M. J., & Russell, R. B. (2003). *Amino Acid Properties and Consequences of Substitutions* (**M. R. Barnes** & **I. C. Gray** Eds.).

Carvill, G. L., McMahon, J. M., Schneider, A., Zemel, M., Myers, C. T., Saykally, J., . . . Mefford, H. C. (2015). Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. *Am J Hum Genet, 96*(5), 808-815. doi:10.1016/j.ajhg.2015.02.016

Chao, H. T., Chen, H., Samaco, R. C., Xue, M., Chahrour, M., Yoo, J., . . . Zoghbi, H. Y. (2010). Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature, 468*(7321), 263-269. doi:10.1038/nature09582

Chateauvieux, S., Morceau, F., Dicato, M., & Diederich, M. (2010). Molecular and therapeutic potential and toxicity of valproic acid. *J Biomed Biotechnol, 2010*. doi:10.1155/2010/479364

Islam, M. P., Herman, G. E., & de Los Reyes, E. C. (2018). Language Regression in an Atypical SLC6A1 Mutation. *Semin Pediatr Neurol, 26*, 25-27. doi:10.1016/j.spen.2018.04.001

Ittisoponpisan, S., Islam, S. A., Khanna, T., Alhuzimi, E., David, A., & Sternberg, M. J. E. (2019). Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? *J Mol Biol, 431*(11), 2197-2212. doi:10.1016/j.jmb.2019.04.009

Johannesen, K. M., Gardella, E., Linnankivi, T., Courage, C., de Saint Martin, A., Lehesjoki, A. E., . . . Moller, R. S. (2018). Defining the phenotypic spectrum of SLC6A1 mutations. *Epilepsia, 59*(2), 389-402. doi:10.1111/epi.13986

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc, 10*(6), 845-858. doi:10.1038/nprot.2015.053

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2019). VarSome: the human genomic variant search engine. *Bioinformatics, 35*(11), 1978-1980. doi:10.1093/bioinformatics/bty897

Liyanage, V. R., & Rastegar, M. (2014). Rett syndrome and MeCP2. *Neuromolecular Med, 16*(2), 231-264. doi:10.1007/s12017-014-8295-9

Mattison, K. A., Butler, K. M., Inglis, G. A. S., Dayan, O., Boussidan, H., Bhambhani, V., . . . Escayg, A. (2018). SLC6A1 variants identified in epilepsy patients reduce gamma-aminobutyric acid transport. *Epilepsia, 59*(9), e135-e141. doi:10.1111/epi.14531

Meng, X., Wang, W., Lu, H., He, L. J., Chen, W., Chao, E. S., . . . Zoghbi, H. Y. (2016). Manipulations of MeCP2 in glutamatergic neurons highlight their contributions to Rett and other neurological disorders. *Elife, 5*. doi:10.7554/eLife.14199

Neul, J. L., Kaufmann, W. E., Glaze, D. G., Christodoulou, J., Clarke, A. J., Bahi-Buisson, N., . . . RettSearch, C. (2010). Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol, 68*(6), 944-950. doi:10.1002/ana.22124

Neul, J. L., Lane, J. B., Lee, H. S., Geerts, S., Barrish, J. O., Annese, F., . . . Percy, A. K. (2014). Developmental delay in Rett syndrome: data from the natural history study. *J Neurodev Disord, 6*(1), 20. doi:10.1186/1866-1955-6-20

Posar, A., & Visconti, P. (2019). Mild Phenotype Associated with SLC6A1 Gene Mutation: A Case Report with Literature Review. *J Pediatr Neurosci, 14*(2), 100-102. doi:10.4103/jpn.JPN_2_19

Vidal, S., Xiol, C., Pascual-Alonso, A., O'Callaghan, M., Pineda, M., & Armstrong, J. (2019). Genetic Landscape of Rett Syndrome Spectrum: Improvements and Challenges. *Int J Mol Sci, 20*(16). doi:10.3390/ijms20163925

Weaving, L. S., Ellaway, C. J., Gecz, J., & Christodoulou, J. (2005). Rett syndrome: clinical review and genetic update. *J Med Genet, 42*(1), 1-7. doi:10.1136/jmg.2004.027730

Zafar, S., & Jabeen, I. (2018). Structure, Function, and Modulation of gamma-Aminobutyric Acid Transporter 1 (GAT1) in Neurological Disorders: A Pharmacoinformatic Prospective. *Front Chem, 6*, 397. doi:10.3389/fchem.2018.00397

DISCUSIÓN

El RTT es un trastorno del neurodesarrollo que se caracteriza por una desaceleración global del desarrollo psicomotor y una pérdida de las adquisiciones cognitivas adquiridas, que ocurre principalmente en niñas después de 6 a 18 meses de desarrollo aparentemente normal, siendo la segunda causa de discapacidad intelectual en mujeres, después del síndrome de Down. Actualmente, cuando han transcurrido más de cincuenta años de la primera publicación donde se definió el RTT a nivel clínico, y poco menos de dos décadas después del primer informe que vincula el RTT con el gen *MECP2*, el esfuerzo de la comunidad científica enfocada a este trastorno se ha centrado en obtener una mejor comprensión de la compleja biología de este trastorno.

En el entorno sanitario, la prioridad que nos planteamos es la mejora diagnóstica tanto a nivel clínico como genético. La recopilación de datos clínicos ha sido imprescindible para poder realizar un análisis de los datos genómicos de las pacientes de manera concisa y poder alcanzar el objetivo principal de la presente tesis, la mejora del diagnóstico genético y la posibilidad de dar un pronóstico sobre el desarrollo de la enfermedad en pacientes con RTT/RTT-*like,* así como profundizar en el estudio de pacientes sin diagnóstico genético para poder determinar la etiología de su clínica. Hemos alcanzado este objetivo trabajando en cooperación con un gran grupo de médicos y hospitales de toda España y del extranjero, para obtener todas las muestras y datos necesarios de pacientes con RTT. Asimismo, los estudios realizados en la presente tesis han servido para dar un resultado genético a pacientes diagnosticadas clínicamente con RTT y que carecían de diagnóstico genético hasta el momento, aumentando de este modo el *ratio* de casos con diagnóstico genético positivo.

Como reflexión general, uno de los hallazgos iniciales que cambiaron el planteamiento inicial de los estudios que se empezaron a realizar utilizando la NGS: ya no esperamos encontrar un único gen para esos casos de RTT sin diagnóstico en los genes *MECP2*, *CDKL5* y *FOXG1*. Sino que existe un gran abanico dentro de las vías alteradas en estos genes, en especial en *MECP2*, que su mal funcionamiento puede estar desarrollando en nuestras pacientes un fenotipo solapante con el RTT.

1. Rendimiento diagnóstico de las nuevas tecnologías

En los últimos años, varios genes se han asociado con fenotipos similares a RTT (Sajan et al., 2017, Olson et al., 2015, Gold and Christodoulou, 2015, Lopes et al., 2016) debido a la incorporación de la NGS. Los métodos de detección tradicionales para genes individuales como la secuenciación Sanger proporcionan un espectro limitado de mutaciones que podrían causar la enfermedad, consumiendo una gran cantidad de tiempo durante su proceso. La NGS, que ha revolucionado la investigación en genética molecular, es un método de alto rendimiento capaz de secuenciar rápidamente una gran cantidad de genes en paralelo y proporcionar grandes conjuntos de datos (DePristo et al., 2011, Marx, 2013).

El objetivo principal de la presente tesis consiste en la mejora del diagnóstico genético, por lo que realizamos un análisis retrospectivo de los distintos métodos utilizados para el estudio molecular de las pacientes con RTT en el laboratorio para determinar cuál de ellos es el procedimiento y la metodología más eficiente, a nivel coste-rendimiento. Desde los inicios, donde el procedimiento se centraba en el estudio mediante secuenciación Sanger para mutaciones puntuales y MLPA para las alteraciones en el número de copias, hasta ahora, donde el método principal es la NGS tanto para mutaciones puntuales como grandes reordenamientos. En el segundo trabajo analizamos los resultados de más de 1000 pacientes estudiadas en nuestro laboratorio durante la última década, en el cual se ha usado distintas metodologías para su análisis molecular:

- Secuenciación Sanger.
- Panel prediseñado de 17 genes asociado con fenotipos RTT y RTT-*like*.
- Exoma clínico en el cual se secuencian 4813 genes asociados a enfermedades ya descritas.
- WES, exoma completo.

En primer lugar, realizamos un análisis y una comparativa de la calidad de los parámetros de las metodologías técnicas empleadas ya que son parte esencial para poder aplicarlas en el diagnóstico genético. La cobertura es superior en los paneles NGS en comparación con WES cuando la cantidad de nucleótidos secuenciados es la misma, en nuestro panel la cobertura es ostentosamente superior a otros métodos de captura, con regiones cubiertas a más del 500x. Comúnmente, la secuenciación de Sanger todavía se aplica a regiones que son

"complejas" para la tecnología NGS complementando así, en los paneles de genes, las regiones mal cubiertas. Estas limitaciones pueden ser técnicas o biológicas, donde se incluyen las regiones ricas en GC y las regiones de alta homología como zonas repetitivas o con pseudogenes. Aunque a diferencia de los paneles NGS, la adición de la secuenciación de Sanger para regiones con cobertura inadecuada no sería rentable para los WES, donde se están analizando más de 20,000 genes en un único individuo (LaDuca et al., 2017). Por el contrario, las estimaciones actuales marcan que la cobertura mínima media que alcanza la secuenciación completa del exoma es de 100x y entre el 90 y 95% de las bases secuenciadas llegan al menos al 10x de cobertura (Rehm et al., 2013).

Por el contrario, un inconveniente clave de los paneles específicos es que pueden quedar obsoletos con bastante rapidez. Durante el tiempo en que se desarrolla y valida un panel para su utilización clínica, pueden haber sido publicados nuevos estudios que identifican y sugieren nuevos genes para enfermedades recientemente caracterizadas. Como en el caso de nuestro panel prediseñado con los 17 genes asociados a fenotipos RTT-*like*, desarrollado en 2012, y que en la actualidad ha quedado obsoleto ya que otros genes como *WDR45, SYNGAP1, IQSEC2, GRIN2B y GABBR2*, han sido descritos en pacientes con fenotipo RTT-*like* con mayor frecuencia que algunos de los genes que ya se encontraban en el panel (Olson et al., 2015, Hoffjan et al., 2016, Yoo et al., 2017, Sajan et al., 2017). De hecho, entre los resultados positivos de WES, el 23% pertenecen a genes caracterizados en los últimos dos años y el 7% son descubrimientos de genes nuevos (Farwell et al., 2015). Como tal, una ventaja de WES es la capacidad de secuenciar todo el exoma a la vez, lo que permite el análisis e interpretación de todas las alteraciones en genes bien caracterizados y recientes poco descritos todavía, y también permite la reinterpretación a medida que se establecen nuevas asociaciones genéticas a lo largo del tiempo. Las ventajas adicionales de un enfoque de secuenciación WES incluye la capacidad de analizar un número significativamente mayor de genes a un coste razonable, el potencial de identificar genes nuevos y la capacidad de secuenciar los exomas de múltiples miembros de la familia simultáneamente en laboratorios que ofrecen secuenciación en tríos y poder establecer los patrones de herencia de las variantes obtenidas *in situ*.

Otra parte importante en el uso de la NGS es la posibilidad de analizar las CNVs utilizando la misma tecnología. Aunque en la actualidad el *gold standard* para el estudio de grandes deleciones y duplicaciones es el MLPA cuando queremos estudiar pocos genes o un

único gen o el CGHa para el estudio de los grandes reordenamientos a nivel genómico, la NGS se ha convertido en una estrategia popular para la caracterización integral de las CNV al generar cientos de millones de lecturas cortas en una sola ejecución. En comparación con los enfoques basados en matrices, donde las sondas están predefinidas para regiones genómicas limitadas, las lecturas cortas de las plataformas NGS se toman al azar de todo el genoma. Las ventajas del enfoque NGS también incluyen una mayor cobertura y resolución, una estimación más precisa de los números de copias, una detección más precisa de los puntos de corte y una mayor capacidad para identificar CNV (Meyerson et al., 2010, Alkan et al., 2011). Teniendo en cuenta estas ventajas, se ha desarrollado un conjunto diverso de herramientas para detectar CNV en función de diferentes características que se pueden extraer de los datos de NGS (Zhao et al., 2013).

Por ello analizamos los datos de las coberturas obtenidas por las distintas metodologías empradas. De este modo hemos conseguido detectar tres CNVs que resultaron ser *de novo* en nuestras pacientes: una deleción de los genes contiguos *IQSEC2* y *KDM5*, y otra de más de 5'7Mb en el cromosoma 15 (15q11.2-15q12). Y una duplicación de 16Mb en el brazo largo del cromosoma 14 (14q32.12-14q32.33). Sin embargo, toda posible CNV detectada mediante NGS ha de ser confirmada por otra técnica complementaria, ya sea MLPA o CGHa u otras técnicas más específicas como la PCR cuantitativa a tiempo real o la PCR digital.

Cuanto más pequeño es el set de genes que se secuencia, mayores son las coberturas que se consiguen por lo que nos aseguramos de que estamos analizando todas las regiones que captura el panel. Por el contrario, con un mayor número de genes analizados mejoramos el *ratio* de diagnóstico y unificamos proceso en el laboratorio ya que en un mismo ensayo podemos agrupar muestras de diversos pacientes con fenotipos completamente diferentes abaratando costes y disminuyendo tiempo de ejecución. Por otro lado, para los casos más complejos en los que no se da con un resultado positivo en genes ya descritos, entramos en el campo de la investigación donde la mejor opción es realizar la secuenciación completa del exoma del caso índice junto con el de los progenitores (WES trio) que nos proporciona de forma inmediata información a nivel de herencia y nos permite centrarnos en aquellas variantes que sean *de novo* en el *probandus* o buscar posibles herencias recesivas detectando variantes en *trans* o en homocigosis. Asimismo, nos permite detectar nuevos genes candidatos que puedan estar asociados con la enfermedad que presenta la paciente y que en la actualidad aún no han sido relacionados, como en nuestro caso en las pacientes con

mutaciones en los genes asociados a la vía gabaérgica (*SLC6A1* y *GABBR2*). En nuestro primer trabajo donde se analizó por WES trio a una cohorte de pacientes con clínica RTT o RTT-*like* obtenemos resultados favorables en un 66% de las pacientes, aproximadamente los mismos porcentajes que se han obtenido en otros laboratorios donde se han realizados estudios similares (Iwama et al., 2019, Lopes et al., 2016, Sajan et al., 2017). Sin embargo, la mayoría de estos casos provienen de previos análisis más exhaustivos negativos, por lo que los resultados que se obtienen acostumbran a necesitar más trabajo de fondo antes de poder dar un diagnóstico a las familias. En la mayoría de casos encontramos mutaciones probablemente patogénicas en genes apenas o nada descritos de los cuales harían falta estudios funcionales complementarios a nivel de gen, para averiguar su función en el organismo, y en la variante encontrada en sí misma, para saber con certeza si realmente ese cambio está produciendo un funcionamiento erróneo de la proteína.

De este modo la secuenciación Sanger en términos de coste-eficiencia ha quedado desbancada a nivel de cribado y diagnóstico molecular. Sin embargo, este método todavía se usa como confirmación de la técnica de NGS y para realizar estudios posteriores de segregación. De este modo, Sanger queda reservado para el estudio de variaciones puntuales que se quieran estudiar en familias donde ya se ha detectado previamente la posible causa de la enfermedad en ella.

Por último, la secuenciación del genoma completo se considera la prueba genética más completa disponible, pero todavía no se aplica en el diagnóstico de rutina de pacientes debido al complejo análisis y manejo de datos que conlleva y el alto coste en comparación con paneles específicos y el WES. Con la secuenciación de todo el genoma podemos detectar, además de SNV en regiones codificantes como en los panes y el WES, variaciones en regiones no codificantes que puedan estar afectando al desarrollo de la enfermedad en el paciente, como pueden ser las regiones reguladoras o mutaciones intrónicas profundas que puedan estar afectando a la transcripción del gen. Sin embargo, en la actualidad no es sostenible el uso de esta metodología para el diagnóstico molecular a nivel asistencial, ya que hacen falta mejoras en el manejo y almacenamiento de datos y que el coste de la secuenciación del genoma completo sea más asequible. Hasta la fecha, no hay publicaciones sobre WGS en cohortes de RTT ni RTT-*like*, pero ya se han publicado estudios asociados a discapacidad intelectual (ID) u otras enfermedades con buenos resultados (Lupski et al., 2010, Gilissen et al., 2014).

2. Correlación genotipo-fenotipo asociado a grandes deleciones en *MECP2*.

El gen *MECP2* se ha asociado con múltiples trastornos neuropsiquiátricos, de los cuales el trastorno del neurodesarrollo RTT es más prominente. Y, aunque han sido realizados múltiples estudios acerca de la correlación genotipo-fenotipo, no ha sido posible en la mayoría de casos establecer una correlación significativa entre las cohortes de pacientes estudiadas. Sin embargo, sí que se ha podido establecer que variantes que truncan la proteína, ya sean variantes *nonsense* como deleciones e inserciones que rompan la pauta de lectura creando un efecto deletéreo, causan un fenotipo más grave en las pacientes. De este modo, en nuestro tercer trabajo hemos caracterizado a nivel molecular las deleciones encontradas en pacientes con fenotipo RTT en nuestro laboratorio acotándolas e intentando detectar el punto de corte de cada una de ellas para poder estudiar los posibles mecanismos que conllevan a este reordenamiento y, además, intentar realizar una correlación genotipo-fenotipo de estas.

Nuestros resultados a nivel molecular mostraron una amplia gama de genotipos, desde deleciones que afectan solamente a una región de un único exón hasta otros que involucran al gen *MECP2* casi en su totalidad y al gen colindante *downstream*, *IRAK1*. Encontramos solamente una paciente con una deleción en los exones 1 y 2 y parte de la región promotora del gen *MECP2*, datos similares a los hallazgos descritos en la literatura, ya que se han sido informadas muy pocas deleciones que afecten a los dos primeros exones del gen (Erlandson et al., 2003, Archer et al., 2006, Hardwick et al., 2007). Cabe destacar, que la mayoría de las deleciones caracterizadas, el punto de corte se encuentra localizado en la región denominada "región propensa a la deleción" localizada en el exón 4 del gen (DPR, *Deletion Prone Region;* GRCh38/hg38 chrX: 154,030,619 - 154,030,770) definido por Laccone en 2004). Además, en esta región también es común encontrar deleciones de tamaños más reducidos (<500 pb). La naturaleza repetitiva de la región DPR, se ha considerado la principal causa de inestabilidad genómica para la creación de deleciones en *MECP2* ya que contiene pequeñas repeticiones directas e invertidas, abundancia de residuos de polipurina en la cadena antisentido y la presencia de la secuencia χ (GCGCTGGTGG), que ha sido descrita como altamente recombinógena en el genoma de la *Escherichia coli* (Stahl, Kobayashi, Stahl, Y Huntington, 1983). Por consiguiente, se ha sugerido que esta secuencia estimula el sistema

dependiente de recombinasa BC y es responsable de ciertas deleciones que causan enfermedades en humanos (Marshall et al., 1996).

A nivel clínico, hemos intentado establecer una correlación genotipo-fenotipo con nuestra cohorte de pacientes que presentan una deleción en *MECP2* y su gravedad clínica utilizando el *score* diseñado por la doctora Mercè Pineda para evaluar la gravedad del fenotipo de las pacientes (Monros et al., 2001). Y, aunque no ha surgido una correlación estadísticamente significativa entre los exones delecionados y la gravedad clínica de este estudio, podemos apreciar algunas tendencias. El paciente descrito con la deleción que involucra a la región promotora de *MECP2* junto con los exones 1 y 2 presenta un fenotipo más severo. Por el contrario, los pacientes descritos con una deleción parcial en el exón 4 muestran el fenotipo más leve en comparación con los demás grupos de paciente, ya que estas conservan los dominios principales de la proteína.

Finalmente, es necesario remarcar las limitaciones que conllevan los estudios realizados de correlaciones genotipo-fenotipo en enfermedades minoritarias como el RTT. Al tratarse de enfermedades con una incidencia baja, en la mayoría de casos la cantidad de individuos que se pueden analizar es pequeña y es complicado obtener valores significativos en su análisis. Otra limitación con la que nos encontramos es la edad de las pacientes que forman la cohorte, ya que los estudios clínicos para generar el *score* que nos indica la gravedad de la enfermedad están realizados sobre una edad determinada. De este modo, la puntuación obtenida puede variar en el tiempo ya que a lo largo de la enfermedad pueden ir apareciendo o mejorando síntomas en las pacientes. Sin embargo, estos estudios nos pueden sugerir una tendencia y ayudar a predecir el desarrollo de la enfermedad dependiendo de qué tipo de anomalía genómica presentan.

3. Identificación de nuevos genes asociados con el RTT

En la búsqueda de un nuevo gen que nos explique la etiología de las pacientes sin mutaciones en los tres genes asociados al RTT hasta el momento, nos encontramos, no con un único nuevo gen candidato que explique la clínica, sino con un gran abanico de genes diversos y todos ellos buenos candidatos. Estos nuevos genes han resultado ser tanto genes ya conocidos y bien caracterizados para otras enfermedades minoritarias como genes poco estudiados, si bien con una posible implicación en el desarrollo y el correcto funcionamiento del SNC. En solamente cinco años más de ochenta genes se relacionaron con este subgrupo de pacientes RTT-*like* sin diagnóstico genético.

3.1 Genes ya descritos para otras patologías

En nuestra primera cohorte de pacientes analizadas por WES y descritas en el primer trabajo hemos podido observar que algunas de ellas presentan variantes patogénicas o probablemente patogénicas en genes ya descritos para otras enfermedades pero que presentan un cuadro clínico solapante con el RTT y muchas veces difícilmente distinguibles entre si dependiendo en la etapa en que se encuentre la paciente.

En este grupo de genes ya asociados a enfermedades conocidas como es el caso del gen *TCF4* (*transcription factor 4*; OMIM*602272), asociado al síndrome de Pitt-Hopkins (PTHS; OMIM#610954), trastorno caracterizado por presentar una discapacidad intelectual severa, un retraso en el desarrollo motor, epilepsia, hipotonía, microcefalia, hiperventilación intermitente y los movimientos estereotipados, junto con unos rasgos faciales característicos y muy comunes en este síndrome como la boca y nariz anchas con puente nasal elevado y una curva pronunciada en el labio superior. Estos rasgos faciales distintivos son más consistentes con el PTHS y lo ayudan a distinguir del RTT, ya que el resto de sintomatología es solapante, pero no todos los pacientes presentan claramente estas dismorfias o a menudo no están todavía bien definidas durante los primeros años de vida. Esto ha llevado a que no solamente nosotros hayamos encontrado en pacientes con un fenotipo RTT o RTT-*like* mutaciones asociadas a este gen, sino que en otros grupos donde se ha planteado la misma estrategia para elucidar sus casos no resueltos relacionados con RTT hayan también encontrado variantes patogénicas en este gen (Lopes et al., 2016, Sajan et al., 2017, Iwama et al., 2019, Schonewolf-Greulich et al., 2019, Srivastava et al., 2018). Otro gen al que le ocurre algo similar es el *GRIN2B* (*glutamate receptor, ionotropic, N-methyl-D-aspartate, subunit 2B*;

OMIM*138252), que está asociado a discapacidad intelectual (OMIM#613970) y encefalopatía epiléptica de inicio infantil (OMIM#616139), ambos, trastornos del neurodesarrollo, caracterizados por un retraso del desarrollo psicomotor con discapacidad intelectual y epilepsia de severidad variable. Otras características que pueden presentar pacientes que tengan alterada la función de este gen son hipotonía, movimientos anómalos como la distonía y rasgos del espectro autista. Todo ello también lleva al solapamiento del cuadro clínico con el RTT, donde otros grupos también han encontrado variantes patogénicas en sus cohortes de pacientes (Sajan et al., 2017, Iwama et al., 2019).

Asimismo, en nuestro primer estudio de WES hemos podido encontrar genes también asociados a encefalopatía epiléptica de inicio en la infancia (*HCN1* y *SCN1A*), que, aunque no han sido detectados en otros estudios similares, si que han sido encontrados otros genes también asociados a una patología similar (*STXBP1, SCN2A, SCN8A, GRIN2A, KCNA2, EEF1A2, KCNB1* y *SYNGAP1*) (Olson et al., 2015, Lopes et al., 2016, Sajan et al., 2017, Allou et al., 2017, Iwama et al., 2019, Schonewolf-Greulich et al., 2019). Con todos los datos obtenidos y viendo que existe un solapamiento de clínica para mejorar el diagnóstico genético en el laboratorio se diseña un panel con los genes con clínica solapante que pueden ser buenos candidatos a analizar a parte de los tres genes conocidos y bien establecidos a RTT. De este modo, en nuestro tercer trabajo, damos a conocer los resultados más significativos obtenidos en nuestro laboratorio con el uso de la NGS desde sus inicios hasta la actualidad. Donde podemos ver que en nuestra cohorte destacan 5 genes como los más representados en nuestra cohorte de pacientes, sin tener en cuenta *MECP2, CDKL5* y *FOXG1*:

El más representado es el gen *STXBP1* (*sintaxin-binding protein-1*; OMIM*602926), asociado a encefalopatía epiléptica de inicio en la infancia (OMIM#612164) y a una serie de síndrome con trastorno del neurodesarrollo como el síndrome de Dravet, de Ohtahara y de West (Carvill et al., 2014, Saitsu et al., 2012, Saitsu et al., 2010). En este caso, nuestros pacientes con variantes patogénicas en este gen presentan a parte de un retraso en el desarrollo psicomotor y epilepsia, una combinación de estereotipias, rasgos autistas y regresión que ya se han descrito en pacientes con variantes en este gen (Olson et al., 2015). Asimismo, el siguiente gen alterado con más pacientes RTT-*like* afectos en nuestra cohorte es el gen *TCF4*, el cual ya se ha comentado en el inicio del apartado sus similitudes con el RTT y RTT-*like*.

El tercer gen en la lista es el *SCN2A* (*sodium cannel, voltage-gated, type II, alpha subunit*; OMIM*182390) que está asociado a tres trastornos: las variantes asociadas con un aumento de función conducen a una encefalopatía epiléptica de inicio infantil (OMIM*613721) o a convulsiones benignas en la infancia familiares (OMIM*607745) y las variantes de pérdida de función a una discapacidad intelectual con trastorno del espectro autista (Sanders et al., 2018). Seguidamente se encuentra el gen *MEF2C* (*mads box trasncription enhancer factor 2, polypeptide C*; OMIM*600662) ha sido asociado a la haploinsuficiencia de *MEF2C* (OMIM#613443) que puede llegar a parecerse al RTT especialmente a la variante de epilepsia precoz. Estos pacientes presentan hipotonía, discapacidad intelectual y retraso motor severo, convulsiones de inicio temprano y estereotípias manuales (Zweier et al., 2010) (Zweier et al., 2010). Esta similitud clínica podría explicarse por una vía común subyacente ya que los pacientes con defectos de *MEF2C* también muestran una disminución de la expresión de *MECP2* y *CDKL5*, como lo demuestran Zweier et al. (2010). Finalmtente, tenemos al gen *SYNGAP1* (*synaptic RAS-GTPase-activating protein 1*; OMIM*603384) ha sido asociado a discapacidad intelectual con epilepsia (OMIM#306684), siguiendo el mismo solapamiento clínico, los pacientes con mutaciones en este gen presentan hipotonía, trastorno de la marcha.

Globalmente, cabe destacar que, con respecto al número de pacientes descritos hasta el momento, hay algunos genes que están más representados y qué estos resultados coinciden con los obtenidos en nuestro cuarto trabajo donde ponemos en común los pacientes con los genes más representados en nuestra cohorte. De este modo, unificando todos los resultados que existen por el momento, hemos podido obtener en nuestro quinto trabajo los genes más representados en el espectro del RTT/RTT-like: *STXBP1, TCF4, SCN2A, WDR45, MEF2C, SHANK3, KCNQ2 y SYNGAP1*.

Puede ser razonable alejarse de las definiciones clínicas de trastornos genéticos hacia definiciones moleculares o biológicas de trastornos, especialmente para RTT. Ciertamente, hay beneficios a la hora de definir a los pacientes por descripciones clínicas, como agrupar pacientes con características similares para el manejo clínico. Sin embargo, este enfoque no funciona bien cuando se trata de ensayos clínicos para terapias que se dirigen a los mecanismos de la enfermedad centrados en el defecto genético central. Por ejemplo, un tratamiento que se dirige a defectos de maduración neuronal observados en mutaciones *MECP2* puede no ser efectivo para una canalopatía iónica debido a alteraciones de *KCNB1*. Además, el fenotipado neurológico y conductual exhaustivo puede revelar distinciones claras

y sutiles que confieren especificidades a cada trastorno genético asociado con RTT. Asimismo, cabe destacar que están siendo detectadas actualmente más comúnmente variantes patogénicas en individuos RTT-*like* los genes *WDR45, STXBP1* y *SHANK3* que en los genes *FOXG1* y *CDKL5*, dos genes establecidos que causan RTT atípica.

Finalmente debemos remarcar el hecho de que todos estos estudios publicados presentan ciertas limitaciones. Separadamente, todos ellos presentan un bajo número de casos positivos en la cohorte de estudio haciendo complicado la obtención de un valor estadístico ya que existen pocos pacientes con mutaciones en un mismo gen que puedan ser asociados a un fenotipo RTT-*like*. Asimismo, las cohortes de pacientes en los que se les aplica el estudio por NGS está formada por pacientes a los que se ya tenían estudios previos negativos los tres genes asociados al RTT o en los que el clínico tiene una posible sospecha de que el fenotipo de la paciente pudiera cuadrar con el cuadro clínico asociado al RTT, y no necesariamente porque el paciente cumplía estrictamente con todos los criterios de diagnóstico para RTT, por lo que es complicado tener cohortes grandes de pacientes con un fenotipo estrictamente homogéneo. Sin embargo, este enfoque refleja la posibilidad de la práctica de los neurólogos junto a los genetistas, donde la NGS puede ayudar a establecer a qué enfermedad puede estar asociado el fenotipo del paciente y ayudar a definir mejor la etiología de esta.

3.2 Genes candidatos no asociados previamente a patología

Por el momento, más de 80 genes han sido asociados al RTT/RTT-*like*, pero en la gran mayoría esta asociación ha de ser críticamente evaluada y contrastada. En nuestro quinto trabajo, realizamos una búsqueda exhaustiva de los resultados de NGS publicados hasta el momento en cohortes de pacientes RTT/RTT-*like* donde hemos podido detectar que los nuevos genes candidatos para explicar el fenotipo de las pacientes se pueden asociar en dos grupos principales por las funciones que desempeña: genes involucrados en la modulación de la cromatina como *SMARCA1, NCOR2, ZSCAN12, ANKRD31* y *ZFX* y genes que participan en la función sináptica como *CACNA1I, SCG2* y *ATP6V0A1*. El vínculo de todos estos genes con las mismas vías en las que están involucrados podría explicar por qué los fenotipos de nuestras pacientes se solapan, causando una función sináptica alterada, trastornos del sueño y una desregulación importante de la expresión génica. Por otra parte, también hay algunos genes involucrados en los procesos de ubiquitinación, a los que también

está vinculado el gen *UBE3A,* principal causa en los pacientes con síndrome de Angelman. La considerable superposición de las características de RTT y Angelman podría explicarse debido a esta posible relación.

Un claro ejemplo de como un nuevo gen candidato ha sido estudiado y definitivamente asociado al fenotipo RTT-like es el gen *GABBR2* (*G Protein-coupled receptor 51*; OMIM*607340), involucrado en la función sináptica de la vía GABAérgica. El ácido γ-amino butírico (GABA) es el neurotransmisor inhibidor más importante del SNC y se conoce que la disminución de la función GABAérgica en sólo un 30-40% conduce a la manifestación de alteraciones neuropsiquiátricas y anomalías funcionales, incluyendo la epilepsia (Chao et al., 2010). Por consiguiente, en nuestro último trabajo iniciamos los estudios funcionales de dos variantes candidatas, detectadas en los estudios WES realizados en nuestro primer trabajo, en genes que desempeñan un papel importante en la vía GABAérgica. La variante *de novo,* p.A567T, detectada en el gen *GABBR2* que codifica para la subunidad 2 del receptor de GABA, ha sido descrita como claramente patogénica con un extenso estudio funcional publicado en 2017 (Yoo et al., 2017). Este estudio proporcionó una evidencia directa de que la señalización de GABA mediada por GABABR alterada conduce a un fenotipo RTT-*like*. Asimismo, posteriormente, esta variante ha sido detectada en diferentes individuos, no relacionados, que presentan un fenotipo RTT-*like*, incluida nuestra paciente (Vuillaume et al., 2018, Samanta and Zarate, 2019). Además, Yoo et al., 2017 identificaron otras mutaciones en el mismo gen (p.S695I y p.I705A) ubicadas en TM6, tienen más probabilidades de afectar la integridad estructural de *GABBR2*, mientras que la mutación RTT (p.A567T), ubicada en TM3, estaría involucrada en la vía de activación: objetivo potencial para el diagnóstico y tratamiento tempranos de RTT y pacientes con EE con mejores pronósticos.

Asimismo, siguiendo los mismos pasos que en *GABBR2*, en nuestro último trabajo realizamos los primeros estudios funcionales en el gen *SLC6A1* (OMIM*137165) y un análisis de la clínica asociada. Este gen codifica para el transportador GAT-1, uno de los principales transportadores de GABA en el cerebro y responsable de la recaptación de GABA en la sinapsis (Madsen et al., 2015). Para este gen hay descrito un amplio abanico de pacientes portadores de variantes patogénicas pero que hasta el momento no han sido asociadas claramente a una clínica RTT-*like*. Sin embargo, la clínica descrita hasta el momento es muy heterogénea, ya que han sido descrito individuos con discapacidad intelectual, con epilepsia mioclonica-atonica, retraso en el desarrollo motor, rasgos autistas y movimientos

estereotipados, y en ciertos casos regresión (Carvill et al., 2015, Islam et al., 2018, Johannesen et al., 2018, Mattison et al., 2018), rasgos compatibles con una clínica RTT y muy similares a los que presenta nuestra paciente con la variante en *SLC6A1*.

Con nuestros estudios preliminares acerca de la patogenicidad del cambio utilizando líneas celulares que expresan el transportador tanto en su forma salvaje como con la mutación presente en nuestra paciente, podemos detectar ciertas anomalías en la localización de la proteína. Sin embargo, aunque podemos ver como menor cantidad de proteína llega a la membrana comparado con la forma salvaje, todavía no podemos descartar que esta desempeñe su función correctamente en la membrana. Asimismo, estudios complementarios a nuestro trabajo sugieren que esta variante detectada en *SLC6A1* pueda ser la causante del fenotipo que presentan las pacientes ya que han conseguido corroborar que, aunque la proteína pueda llegar a membrana, el canal transportador de GABA no funciona correctamente.

3.3 "Rett spectrum disorders"

Poniendo en conjunto todas las cohortes de pacientes RTT que han sido analizadas tanto por nosotros como por el resto de comunidad científica, todo ello resumido y agrupado en nuestro quinto trabajo, podemos destacar que todos los individuos resumidos en esta revisión cumplieron los criterios diagnósticos para RTT o RTT-*like*. Sin embargo, la falta de un defecto en *MECP2* subraya la importancia de realizar estudios mediante NGS para obtener un diagnóstico para estas pacientes. El diagnóstico diferencial en el RTT ha evolucionado con el desarrollo de la NGS y muchos pacientes han sido diagnosticados con otros síndromes o variantes en genes recientemente descritos en los que el fenotipo asociado aún no se ha explorado por completo. Muchos trastornos pueden ser causados por múltiples genes, como el síndrome de West o las enfermedades metabólicas como los trastornos congénitos de la glucosilación, y estos se consideran el mismo trastorno siempre que compartan un fenotipo común. Por lo tanto, consideramos que todos los fenotipos clásicos, atípicos y RTT-*like* podrían agruparse en un trastorno del espectro RTT con un gran número de genes causales, aunque deberían ser establecidos y consensuados los criterios para poder definir este grupo de enfermedades.

4. Definiendo las vías alteradas por *MECP2*

Dada la complejidad de los síntomas clínicos, estudiar la relación de los trastornos neurológicos pediátricos, como el RTT, sigue siendo un gran desafío. La gran diversidad de genes con variantes patogénicas o probablemente patogénicas detectados hasta el momento nos indica que todos estos fenotipos solapantes al RTT se encuentran bajo rutas biológicas con una gran complejidad en las que *MECP2* se encuentra integrado. Comprender esta compleja red podría ser importante en el desarrollo de terapias efectivas, dirigidas específicamente a los puntos donde el sistema está fallando.

4.1 Análisis de enriquecimiento

Para delinear las rutas biológicas y las redes de genes que interactúan y están involucrados en fenotipos similares al RTT varios grupos de investigación han intentado realizar asociaciones de los genes detectados hasta el momento y evaluar sus propiedades funcionales en conjunto. Los primeros estudios realizados en pacientes RTT negativos para variantes patogénicas en *MECP2* no generaron resultados claros ni recurrentes que pudieran agrupar los genes detectados hasta el momento o se pudieran detectar vías claras de señalización, aunque si se destacaba un claro enriquecimiento asociado a las vías de señalización de glutamato y genes relacionados con la formación de sinapsis (Olson et al., 2015, Sajan et al., 2017). Sin embargo, recientemente, Iwama et al., 2019 realizaron un análisis de enriquecimiento con la base de datos de *Gene Ontology* (Go) con 50 genes con mutaciones patogénicas o probablemente patogénicas con fenotipos similares al RTT y evaluaron sus propiedades funcionales. El análisis de enriquecimiento de GO de estos 50 genes identificó un enriquecimiento significativo de varios términos GO, como transporte de iones, señalización sináptica y desarrollo del sistema nervioso. Con base a la información de estos términos asignados para cada gen, clasificaron estos 50 genes en cuatro grupos biológicos distintos: (1) señalización sináptica que incluyen receptores de glutamato o GABA y otros reguladores de señalización sináptica (p. Ej., *GRIN2B* e *IRF2BPL*); (2) transporte de iones donde se incluyen canales de sodio, potasio y calcio y transportadores de protones (p. Ej., *CACNA1G* y *ATP6V0A1*); (3) unión al ADN para la regulación de la transcripción (p. Ej., *NR2F1* y *NCOR2*) y (4) otros, incluidos la ruta de la autofagia, la función enzimática y los componentes de la estructura de la matriz (p. Ej., *WDR45* y *PPT1*).

Del mismo modo, en nuestro quinto trabajo, realizamos un análisis de los más de 80 genes que han sido asociados hasta el momento con los fenotipos RTT/RTT-*like* haciendo uso de REACTOME, base de datos de rutas biológicas de acceso libre y curada manualmente, y de STRING, base de datos de interacciones proteína-proteína conocidas y predichas. Con los datos obtenidos, hemos podido identificar algunos grupos de genes con funciones involucradas en mecanismos comunes como en los estudios realizados por Iwama et al., 2019. De este modo, hemos detectado varias vías que parecen ser relevantes en pacientes con RTT / RTT-*like*, como la modulación de la cromatina, la función sináptica y la conjugación de ubiquitina. De este modo podemos ampliar los listados de genes para el análisis de exomas/genomas de nuevas pacientes en las que no se les detecte ninguna variante candidata que explique su fenotipo en los genes principales ya descritos asociados a la clínica RTT/RTT-*like*.

En el enriquecimiento de GO y los análisis interactivos de la red de genes de los relacionados con el RTT, encontramos el transporte de iones y la señalización sináptica como las principales vías biológicas involucradas en los fenotipos similares a RTT. Se sabe que las variantes patogénicas en los genes relacionados con la transducción de señales o el transporte de iones (canales y transportadores) causan retraso en el desarrollo/discapacidad intelectual, epilepsia y otros trastornos del neurodesarrollo. Del mismo modo, era un resultado esperado encontrar variantes patogénicas en genes que codifican proteínas de unión al ADN, como *TCF4* y *DNMT3A*, cuyas anomalías ya se asociaron originalmente a diferentes síndromes con un fenotipo solapante al RTT (p. Ej., Síndrome de Pitt-Hopkins y síndrome de TattonBrown-Rahman, respectivamente). Por otra parte, como se han observado previamente desequilibrios en los sistemas GABAérgico y glutamatérgico en los cerebros de ratones con deficiencia de Mecp2 (El-Khoury et al., 2014), sería razonable observar variantes en estos genes en fenotipos similares a RTT. De ahí que, variantes detectadas en genes relacionados con estos sistemas hayan sido detectadas en pacientes con un fenotipo RTT-*like* y sin un diagnóstico genético positivo en los genes que hasta el momento estaban relacionados con la patología.

Asimismo, estos estudios de enriquecimiento también nos han dado a conocer varios factores de transcripción que interactúan con *MECP2*: *E2F1, MEF2C, REST, SIN3A, SMC3, SP1* y *TAF1*. *E2F1* y *TAF1* son elementos promotores de *MECP2* (Ehrhart et al., 2016). *SP1* se conoce como un elemento regulador *cis* y un elemento promotor de *MECP2*.

REST es un elemento regulador *cis* de *MECP2*. *SIN3A* es una parte integral del complejo *MECP2*-HDAC y *SMC3* es uno de los cofactores de este complejo. La expresión de *MEF2C* es inhibida por *MECP2* junto con el complejo HDAC. Interactúa con *EP300*, varias histonas desacetilasas y *SP1*. Además, estudios de enriquecimiento de vías con factores de transcripción conocidos ha revelado una posible conexión entre la organización del citoesqueleto y *MECP2* a través de *MEF2C* y *CAPG*. Por lo que se presume que *MECP2* regula la expresión de *CAPG* a través de *MEF2C* (Ehrhart et al., 2019).

4.2 Análisis transcriptómicos

Si bien los mecanismos de acción de MeCP2 aún no se han caracterizado por completo, está claro que la proteína tiene un efecto importante en la regulación transcripcional. Centrar la atención en los cambios en la expresión génica que ocurren en individuos con RTT puede ser una forma de dilucidar el vínculo entre la desregulación de genes / vías candidatas y el fenotipo resultante. A pesar de que por el momento existe un número limitado de estudios a nivel transcriptómico en pacientes con RTT, estos han proporcionado información interesante para comprender parte de la fisiopatología de la enfermedad. La principal conclusión que surge de los estudios transcriptómicos es la convergencia de mecanismos a través de diferentes tejidos. Los genes desregulados pertenecen a tres categorías principales: arborización dendrítica anormal y maduración sináptica, disfunción mitocondrial y actividad de células gliales (Shovlin and Tropea, 2018). Asimismo, hay estudios que confirman que las vías inflamatorias, de procesamiento de ARNm y de iniciación de la transcripción se ven afectadas cuando hay una alteración en *MECP2,* e incluso pueden ser los principales impulsores del fenotipo que desarrollan las pacientes con RTT (Ehrhart et al., 2019).

Por otro lado, un trabajo reciente ha sugerido la posibilidad de la participación del ADNmt en el desarrollo del fenotipo RTT (Aldosary et al., 2020), ya que se sabe que en muestras de fibroblastos y músculo de pacientes con RTT presentan anomalías tanto en el número como en el tamaño de las mitocondrias (Cardaioli et al., 1999). Del mismo modo, diferentes estudios de transcriptómica han revelado que existe una disfunción mitocondrial (Shulyakova et al., 2017, Shovlin and Tropea, 2018). Y, recientemente, han sido identificados roles potencialmente importantes de genes involucrados en la disfunción mitocondrial y la

respuesta al estrés oxidativo, incluidos los genes *NDUFA5, MAPK9, ATP5PO, MAF, FOS* y *SAPK,* en la patogenicidad del RTT (Ehrhart et al., 2019).

Finalmente, también se ha podido comprobar que no solamente la expresión de genes está siendo alterada, sino que hay ciertos RNAs no codificantes de cadena larga (lncRNA; *long non-coding RNA*) se encuentran alterados cuando MeCP2 no funciona correctamente. Los estudios de transcriptómica en modelos de ratón nulos para Mecp2 nos han dado a conocer lncRNAs que son regulados por Mecp2: AK081227 y AK087060, ya que sus niveles se ven alterados en comparación con los controles. Es importante destacar que la expresión de AK08127 mediada por la pérdida de Mecp2 inhibió a la subunidad del receptor de ácido gamma-aminobutírico Rho 2 (Gabrr2) que está involucrada en la neurotransmisión GABAérgica (Petazzi et al., 2013). Estos hallazgos resaltan el papel de los lncRNAs en el fenotipo neurológico y la etiología de RTT.

Una conclusión general de estos análisis es que una enfermedad monogénica es definitivamente verdadera como base para la RTT, pero el vínculo entre la mutación y el resultado fenotípico está definido por una gama más amplia de vías y procesos en el desarrollo neurológico. Esta conexión implica una red compleja en la que la variación genética individual y la influencia ambiental determinan el resultado del trastorno. Aunque se requiere un análisis adicional para confirmar los eventos fisiopatológicos exactos que tienen lugar en pacientes con RTT, los estudios transcriptómicos representan una muy buena base imparcial para la detección de comportamientos celulares aberrantes y proporcionan a los investigadores una hoja de ruta para guiar investigaciones específicas. Debido a la amplitud de detección y sensibilidad de estos estudios, sus hallazgos pueden usarse para generar nuevas hipótesis para ser probadas en conjuntos adicionales de experimentos. Dichos análisis pueden usarse para descubrir la base biológica de RTT y señalar nuevas estrategias para intervenciones.

5. Perspectivas diagnósticas y terapéuticas

Como clausura de los trabajos presentados, fomentamos las nuevas estrategias para abordar el diagnóstico genético y poder implantarlo a nivel asistencial. Estos últimos años, el algoritmo diagnóstico ha evolucionado considerablemente, del análisis de un único gen a analizar de decenas a miles a la vez. De este modo, implementando en el laboratorio la secuenciación masiva, aumentamos el rendimiento diagnóstico a nivel molecular dando resultados positivos a un número mayor de familias.

Desde el punto de vista de la investigación, los nuevos genes recientemente asociados a este espectro de enfermedades RTT-*like* sin una descripción clara de las vías en las que se encuentran necesitan ser estudiadas más a fondo. Estos nuevos genes candidatos necesitan estudios funcionales para establecer su papel potencial en la patogénesis de la enfermedad. Además, la aclaración de las vías funcionales involucradas podría allanar el camino para el desarrollo de futuras terapias dirigidas, tanto en pacientes que tengan ese gen en concreto alterado u otros que presenten alterados otros genes dentro de las mismas vías. Asimismo, conociendo cual es el defecto genético que tiene el paciente, se le podrá implementar en un futuro la terapia génica adecuada.

Desde una perspectiva clínica, una mejor definición de las vías que conectan a todos estos genes involucrados en todos los fenotipos RTT y RTT-*like* permitiría una mejor comprensión del panorama genético del espectro RTT. Asimismo, estudiando y comparando las cohortes de pacientes que tengan los mismos genes alterados, a lo largo del desarrollo de la enfermedad, podrá ayudar a pautar cuales pueden ser los mejores tratamientos a seguir. Asimismo, tener clara la etiología de la enfermedad de cada paciente puede ayudar a escoger la terapia farmacológica más adecuada y eficaz como es el caso de las pacientes con defectos en el gen *GRIN2B*. Estudios funcionales recientes de una variante *missense* patogénica en *GRIN2B* demostraron que suplementar la dieta con L-serina de la paciente RTT-*like* que presentaba esta mutación mostro mejoras notables en el rendimiento motor y cognitivo y de la comunicación después de los 11 meses de la suplementación. Sugiriendo que la suplementación con L-serina podría mejorar la encefalopatía grave relacionada con *GRIN2B* y otras afecciones neurológicas causadas por la deficiencia de señalización glutamatérgica (Soto et al., 2019).

La ciencia ha entrado en el mundo de la "omica", donde ya no realizamos análisis de un único gen, sino que somos capaces de estudiar simultáneamente todo el genoma, transcriptoma y proteoma. Y con estos avances podremos diseñar nuevos métodos para la mejora del diagnóstico genético en nuestra cohorte de pacientes que siguen sin un diagnóstico concluyente. En nuestro caso, donde el gen principal de la enfermedad es el *MECP2,* ya que en la mayoría casos presentan este gen alterado, y se ha visto que los casos en los que se ha visto mutado otro gen, este es muy probable que se encuentre dentro de las mismas vías en las que se encuentra o regula *MECP2*. Con esta premisa, podríamos realizar una aproximación *multiómica*, aparte de realizar un análisis a nivel de paneles dirigidos a los genes ya claramente descritos y a un fenotipo dentro del espectro Rett y/o un exoma para buscar posibles nuevos genes candidatos, se podría analizar si los niveles de *MECP2* en estás pacientes sin diagnóstico genético está alterado y a nivel trasncriptómico y/o proteómico qué vías están alteradas

CONCLUSIONES

De los resultados expuestos se puede concluir que:

1. Aunque en la actualidad el diagnóstico del RTT se basa en criterios clínicos, la confirmación a nivel molecular permite un consejo genético y en un futuro poder emplear la terapia génica adecuada.
2. Si bien la mayoría de pacientes RTT son portadoras de variantes patogénicas en el gen *MECP2*, la causa molecular de los pacientes restantes no puede ser atribuida a un solo gen. Se ha conseguido confeccionar un listado de genes candidatos que puedan estar alterando tanto *MECP2* como funciones vitales en el SNC causantes del fenotipo RTT-*like*, habiendo confirmado algunos de ellos.
3. La aproximación metodológica con mejor rendimiento para el estudio molecular del RTT es el estudio del exoma clínico, analizando las regiones codificantes de los genes con descripción fenotípica establecida. Sin embargo, para aquellos casos con resultados negativos a nivel asistencial, la mejor aproximación es el exoma trio.
4. Las grandes deleciones en *MECP2* presentan su punto de rotura en la mayoría de los casos en la región DPR, localizada en el exón 4 del gen. Asimismo, las pacientes con deleciones que afectan a la región promotora del gen o ven afectados sus dominios funcionales tienden a presentar una clínica más severa.
5. Genes asociados a la vía GABAaérgica pueden ser causantes de fenotipos Rett-*like*, tales como el gen *GABBR2* y el gen *SLC6A1*, por lo que genes presentes en esta vía son buenos candidatos para analizar en nuevas pacientes con RTT.
6. Son necesarios estudios funcionales para poder atribuir genes candidatos como causantes de la enfermedad de las pacientes si estos no han sido previamente descritos con anterioridad, más aún si los genes no han sido asociados a alguna enfermedad. Los estudios funcionales realizados en la variante detectada del gen *SLC6A1* sugieren una relación de causalidad entre dicha alteración y el fenotipo de la paciente.

Bibliografía

ALDOSARY, M., AL-BAKHEET, A., AL-DHALAAN, H., ALMASS, R., ALSAGOB, M., AL-YOUNES, B., ALQUAIT, L., MUSTAFA, O. M., BULBUL, M., RAHBEENI, Z., ALFADHEL, M., CHEDRAWI, A., AL-HASSNAN, Z., ALDOSARI, M., AL-ZAIDAN, H., AL-MUHAIZEA, M. A., ALSAYED, M. D., SALIH, M. A., ALSHAMMARI, M., FAIYAZ-UL-HAQUE, M., CHISHTI, M. A., AL-HARAZI, O., AL-ODAIB, A., KAYA, N. & COLAK, D. 2020. Rett Syndrome, a Neurodevelopmental Disorder, Whole-Transcriptome, and Mitochondrial Genome Multiomics Analyses Identify Novel Variations and Disease Pathways. *OMICS,* 24**,** 160-171.

ALKAN, C., COE, B. P. & EICHLER, E. E. 2011. Genome structural variation discovery and genotyping. *Nat Rev Genet,* 12**,** 363-76.

ALLOU, L., JULIA, S., AMSALLEM, D., EL CHEHADEH, S., LAMBERT, L., THEVENON, J., DUFFOURD, Y., SAUNIER, A., BOUQUET, P., PERE, S., MOUSTAINE, A., RUAUD, L., ROTH, V., JONVEAUX, P. & PHILIPPE, C. 2017. Rett-like phenotypes: expanding the genetic heterogeneity to the KCNA2 gene and first familial case of CDKL5-related disease. *Clin Genet,* 91**,** 431-440.

AMARAL, M. D. & POZZO-MILLER, L. 2007. TRPC3 channels are necessary for brain-derived neurotrophic factor to activate a nonselective cationic current and to induce dendritic spine formation. *J Neurosci,* 27**,** 5179-89.

AMIR, R. E., VAN DEN VEYVER, I. B., WAN, M., TRAN, C. Q., FRANCKE, U. & ZOGHBI, H. Y. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet,* 23**,** 185-8.

ANDERSON, A., WONG, K., JACOBY, P., DOWNS, J. & LEONARD, H. 2014. Twenty years of surveillance in Rett syndrome: what does this tell us? *Orphanet J Rare Dis,* 9**,** 87.

ARCHER, H. L., WHATLEY, S. D., EVANS, J. C., RAVINE, D., HUPPKE, P., KERR, A., BUNYAN, D., KERR, B., SWEENEY, E., DAVIES, S. J., REARDON, W., HORN, J., MACDERMOT, K. D., SMITH, R. A., MAGEE, A., DONALDSON, A., CROW, Y., HERMON, G., MIEDZYBRODZKA, Z., COOPER, D. N., LAZAROU, L., BUTLER, R., SAMPSON, J., PILZ, D. T., LACCONE, F. & CLARKE, A. J. 2006. Gross rearrangements of the MECP2 gene are found in both classical and atypical Rett syndrome patients. *J Med Genet,* 43**,** 451-6.

ARIANI, F., HAYEK, G., RONDINELLA, D., ARTUSO, R., MENCARELLI, M. A., SPANHOL-ROSSETO, A., POLLAZZON, M., BUONI, S., SPIGA, O., RICCIARDI, S., MELONI, I., LONGO, I., MARI, F., BROCCOLI, V., ZAPPELLA, M. & RENIERI, A. 2008. FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet,* 83**,** 89-93.

BAHI-BUISSON, N., NECTOUX, J., GIRARD, B., VAN ESCH, H., DE RAVEL, T., BODDAERT, N., PLOUIN, P., RIO, M., FICHOU, Y., CHELLY, J. & BIENVENU, T. 2010. Revisiting the phenotype associated with FOXG1 mutations: two novel cases of congenital Rett variant. *Neurogenetics,* 11**,** 241-9.

BALLESTAR, E., YUSUFZAI, T. M. & WOLFFE, A. P. 2000. Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry,* 39**,** 7100-6.

BEDOGNI, F., COBOLLI GIGLI, C., POZZI, D., ROSSI, R. L., SCARAMUZZA, L., ROSSETTI, G., PAGANI, M., KILSTRUP-NIELSEN, C., MATTEOLI, M. & LANDSBERGER, N. 2016. Defects During Mecp2 Null Embryonic Cortex Development Precede the Onset of Overt Neurological Symptoms. *Cereb Cortex,* 26**,** 2517-2529.

BEYER, K. S., BLASI, F., BACCHELLI, E., KLAUCK, S. M., MAESTRINI, E., POUSTKA, A. & INTERNATIONAL MOLECULAR GENETIC STUDY OF AUTISM, C. 2002. Mutation

analysis of the coding sequence of the MECP2 gene in infantile autism. *Hum Genet,* 111**,** 305-9.

BROWN, K., SELFRIDGE, J., LAGGER, S., CONNELLY, J., DE SOUSA, D., KERR, A., WEBB, S., GUY, J., MERUSI, C., KOERNER, M. V. & BIRD, A. 2016. The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Hum Mol Genet,* 25**,** 558-70.

BRUNETTI-PIERRI, N., PACIORKOWSKI, A. R., CICCONE, R., DELLA MINA, E., BONAGLIA, M. C., BORGATTI, R., SCHAAF, C. P., SUTTON, V. R., XIA, Z., JELLUMA, N., RUIVENKAMP, C., BERTRAND, M., DE RAVEL, T. J., JAYAKAR, P., BELLI, S., ROCCHETTI, K., PANTALEONI, C., D'ARRIGO, S., HUGHES, J., CHEUNG, S. W., ZUFFARDI, O. & STANKIEWICZ, P. 2011. Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment. *Eur J Hum Genet,* 19**,** 102-7.

CARDAIOLI, E., DOTTI, M. T., HAYEK, G., ZAPPELLA, M. & FEDERICO, A. 1999. Studies on mitochondrial pathogenesis of Rett syndrome: ultrastructural data from skin and muscle biopsies and mutational analysis at mtDNA nucleotides 10463 and 2835. *J Submicrosc Cytol Pathol,* 31**,** 301-4.

CARTRON, P. F., NADARADJANE, A., LEPAPE, F., LALIER, L., GARDIE, B. & VALLETTE, F. M. 2013. Identification of TET1 Partners That Control Its DNA-Demethylating Function. *Genes Cancer,* 4**,** 235-41.

CARVILL, G. L., MCMAHON, J. M., SCHNEIDER, A., ZEMEL, M., MYERS, C. T., SAYKALLY, J., NGUYEN, J., ROBBIANO, A., ZARA, F., SPECCHIO, N., MECARELLI, O., SMITH, R. L., LEVENTER, R. J., MOLLER, R. S., NIKANOROVA, M., DIMOVA, P., JORDANOVA, A., PETROU, S., EURO, E. R. E. S. M.-A. E., DRAVET WORKING, G., HELBIG, I., STRIANO, P., WECKHUYSEN, S., BERKOVIC, S. F., SCHEFFER, I. E. & MEFFORD, H. C. 2015. Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. *Am J Hum Genet,* 96**,** 808-15.

CARVILL, G. L., WECKHUYSEN, S., MCMAHON, J. M., HARTMANN, C., MOLLER, R. S., HJALGRIM, H., COOK, J., GERAGHTY, E., O'ROAK, B. J., PETROU, S., CLARKE, A., GILL, D., SADLEIR, L. G., MUHLE, H., VON SPICZAK, S., NIKANOROVA, M., HODGSON, B. L., GAZINA, E. V., SULS, A., SHENDURE, J., DIBBENS, L. M., DE JONGHE, P., HELBIG, I., BERKOVIC, S. F., SCHEFFER, I. E. & MEFFORD, H. C. 2014. GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. *Neurology,* 82**,** 1245-53.

CHAHROUR, M., JUNG, S. Y., SHAW, C., ZHOU, X., WONG, S. T., QIN, J. & ZOGHBI, H. Y. 2008. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science,* 320**,** 1224-9.

CHAHROUR, M. & ZOGHBI, H. Y. 2007. The story of Rett syndrome: from clinic to neurobiology. *Neuron,* 56**,** 422-37.

CHANG, Q., KHARE, G., DANI, V., NELSON, S. & JAENISCH, R. 2006. The disease progression of Mecp2 mutant mice is affected by the level of BDNF expression. *Neuron,* 49**,** 341-8.

CHAO, H. T., CHEN, H., SAMACO, R. C., XUE, M., CHAHROUR, M., YOO, J., NEUL, J. L., GONG, S., LU, H. C., HEINTZ, N., EKKER, M., RUBENSTEIN, J. L., NOEBELS, J. L., ROSENMUND, C. & ZOGHBI, H. Y. 2010. Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. *Nature,* 468**,** 263-9.

CHAPLEAU, C. A., CALFA, G. D., LANE, M. C., ALBERTSON, A. J., LARIMORE, J. L., KUDO, S., ARMSTRONG, D. L., PERCY, A. K. & POZZO-MILLER, L. 2009. Dendritic spine pathologies in hippocampal pyramidal neurons from Rett syndrome brain and after expression of Rett-associated MECP2 mutations. *Neurobiol Dis,* 35**,** 219-33.

CHEN, R. Z., AKBARIAN, S., TUDOR, M. & JAENISCH, R. 2001. Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat Genet,* 27**,** 327-31.

CHEN, W. G., CHANG, Q., LIN, Y., MEISSNER, A., WEST, A. E., GRIFFITH, E. C., JAENISCH, R. & GREENBERG, M. E. 2003. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. *Science,* 302**,** 885-9.

CHEN, Y., YU, J., NIU, Y., QIN, D., LIU, H., LI, G., HU, Y., WANG, J., LU, Y., KANG, Y., JIANG, Y., WU, K., LI, S., WEI, J., HE, J., WANG, J., LIU, X., LUO, Y., SI, C., BAI, R., ZHANG, K., LIU, J., HUANG, S., CHEN, Z., WANG, S., CHEN, X., BAO, X., ZHANG, Q., LI, F., GENG, R., LIANG, A., SHEN, D., JIANG, T., HU, X., MA, Y., JI, W. & SUN, Y. E. 2017. Modeling Rett Syndrome Using TALEN-Edited MECP2 Mutant Cynomolgus Monkeys. *Cell,* 169**,** 945-955 e10.

CHRISTODOULOU, J., GRIMM, A., MAHER, T. & BENNETTS, B. 2003. RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat,* 21**,** 466-72.

COHEN, D., LAZAR, G., COUVERT, P., DESPORTES, V., LIPPE, D., MAZET, P. & HERON, D. 2002. MECP2 mutation in a boy with language disorder and schizophrenia. *Am J Psychiatry,* 159**,** 148-9.

CONRAD, D. F., KEEBLER, J. E., DEPRISTO, M. A., LINDSAY, S. J., ZHANG, Y., CASALS, F., IDAGHDOUR, Y., HARTL, C. L., TORROJA, C., GARIMELLA, K. V., ZILVERSMIT, M., CARTWRIGHT, R., ROULEAU, G. A., DALY, M., STONE, E. A., HURLES, M. E., AWADALLA, P. & GENOMES, P. 2011. Variation in genome-wide mutation rates within and between human families. *Nat Genet,* 43**,** 712-4.

CUDDAPAH, V. A., PILLAI, R. B., SHEKAR, K. V., LANE, J. B., MOTIL, K. J., SKINNER, S. A., TARQUINIO, D. C., GLAZE, D. G., MCGWIN, G., KAUFMANN, W. E., PERCY, A. K., NEUL, J. L. & OLSEN, M. L. 2014. Methyl-CpG-binding protein 2 (MECP2) mutation type is associated with disease severity in Rett syndrome. *J Med Genet,* 51**,** 152-8.

DEACON, R. M., GLASS, L., SNAPE, M., HURLEY, M. J., ALTIMIRAS, F. J., BIEKOFSKY, R. R. & COGRAM, P. 2015. NNZ-2566, a novel analog of (1-3) IGF-1, as a potential therapeutic agent for fragile X syndrome. *Neuromolecular Med,* 17**,** 71-82.

DEATON, A. M., WEBB, S., KERR, A. R., ILLINGWORTH, R. S., GUY, J., ANDREWS, R. & BIRD, A. 2011. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res,* 21**,** 1074-86.

DEOGRACIAS, R., YAZDANI, M., DEKKERS, M. P., GUY, J., IONESCU, M. C., VOGT, K. E. & BARDE, Y. A. 2012. Fingolimod, a sphingosine-1 phosphate receptor modulator, increases BDNF levels and improves symptoms of a mouse model of Rett syndrome. *Proc Natl Acad Sci U S A,* 109**,** 14230-5.

DEPRISTO, M. A., BANKS, E., POPLIN, R., GARIMELLA, K. V., MAGUIRE, J. R., HARTL, C., PHILIPPAKIS, A. A., DEL ANGEL, G., RIVAS, M. A., HANNA, M., MCKENNA, A., FENNELL, T. J., KERNYTSKY, A. M., SIVACHENKO, A. Y., CIBULSKIS, K., GABRIEL, S. B., ALTSHULER, D. & DALY, M. J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet,* 43**,** 491-8.

DI RESTA, C., GALBIATI, S., CARRERA, P. & FERRARI, M. 2018. Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. *EJIFCC,* 29**,** 4-14.

DORE, A. S., OKRASA, K., PATEL, J. C., SERRANO-VEGA, M., BENNETT, K., COOKE, R. M., ERREY, J. C., JAZAYERI, A., KHAN, S., TEHAN, B., WEIR, M., WIGGIN, G. R. & MARSHALL, F. H. 2014. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature,* 511**,** 557-62.

EHRHART, F., COORT, S. L., CIRILLO, E., SMEETS, E., EVELO, C. T. & CURFS, L. 2016. New insights in Rett syndrome using pathway analysis for transcriptomics data. *Wien Med Wochenschr,* 166**,** 346-52.

EHRHART, F., COORT, S. L., EIJSSEN, L., CIRILLO, E., SMEETS, E. E., BAHRAM SANGANI, N., EVELO, C. T. & CURFS, L. M. G. 2019. Integrated analysis of human transcriptome data for Rett syndrome finds a network of involved genes. *World J Biol Psychiatry***,** 1-23.

EHRHART, F., SANGANI, N. B. & CURFS, L. M. G. 2018. Current developments in the genetics of Rett and Rett-like syndrome. *Curr Opin Psychiatry,* 31**,** 103-108.

EINSPIELER, C., KERR, A. M. & PRECHTL, H. F. 2005. Abnormal general movements in girls with Rett disorder: the first four months of life. *Brain Dev,* 27 Suppl 1**,** S8-S13.

EL-KHOURY, R., PANAYOTIS, N., MATAGNE, V., GHATA, A., VILLARD, L. & ROUX, J. C. 2014. GABA and glutamate pathways are spatially and developmentally affected in the brain of Mecp2-deficient mice. *PLoS One,* 9**,** e92169.

ELEFANT, C. & WIGRAM, T. 2005. Learning ability in children with Rett syndrome. *Brain Dev,* 27 Suppl 1**,** S97-S101.

ERLANDSON, A., SAMUELSSON, L., HAGBERG, B., KYLLERMAN, M., VUJIC, M. & WAHLSTROM, J. 2003. Multiplex ligation-dependent probe amplification (MLPA) detects large deletions in the MECP2 gene of Swedish Rett syndrome patients. *Genet Test,* 7**,** 329-32.

EYRE-WALKER, A. & KEIGHTLEY, P. D. 2007. The distribution of fitness effects of new mutations. *Nat Rev Genet,* 8**,** 610-8.

FARWELL, K. D., SHAHMIRZADI, L., EL-KHECHEN, D., POWIS, Z., CHAO, E. C., TIPPIN DAVIS, B., BAXTER, R. M., ZENG, W., MROSKE, C., PARRA, M. C., GANDOMI, S. K., LU, I., LI, X., LU, H., LU, H. M., SALVADOR, D., RUBLE, D., LAO, M., FISCHBACH, S., WEN, J., LEE, S., ELLIOTT, A., DUNLOP, C. L. & TANG, S. 2015. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med,* 17**,** 578-86.

FENG, B. J. 2017. PERCH: A Unified Framework for Disease Gene Prioritization. *Hum Mutat,* 38**,** 243-251.

GENOMES PROJECT, C., ABECASIS, G. R., AUTON, A., BROOKS, L. D., DEPRISTO, M. A., DURBIN, R. M., HANDSAKER, R. E., KANG, H. M., MARTH, G. T. & MCVEAN, G. A. 2012. An integrated map of genetic variation from 1,092 human genomes. *Nature,* 491**,** 56-65.

GILISSEN, C., HEHIR-KWA, J. Y., THUNG, D. T., VAN DE VORST, M., VAN BON, B. W., WILLEMSEN, M. H., KWINT, M., JANSSEN, I. M., HOISCHEN, A., SCHENCK, A., LEACH, R., KLEIN, R., TEARLE, R., BO, T., PFUNDT, R., YNTEMA, H. G., DE VRIES, B. B., KLEEFSTRA, T., BRUNNER, H. G., VISSERS, L. E. & VELTMAN, J. A. 2014. Genome sequencing identifies major causes of severe intellectual disability. *Nature,* 511**,** 344-7.

GOLD, W. A. & CHRISTODOULOU, J. 2015. The Utility of Next-Generation Sequencing in Gene Discovery for Mutation-Negative Patients with Rett Syndrome. *Front Cell Neurosci,* 9**,** 266.

GOS, M. 2013. Epigenetic mechanisms of gene expression regulation in neurological diseases. *Acta Neurobiol Exp (Wars),* 73**,** 19-37.

GUERRINI, R. & PARRINI, E. 2012. Epilepsy in Rett syndrome, and CDKL5- and FOXG1-gene-related encephalopathies. *Epilepsia,* 53**,** 2067-78.

GUIDERI, F. & ACAMPA, M. 2005. Sudden death and cardiac arrhythmias in Rett syndrome. *Pediatr Cardiol,* 26**,** 111.

GULMEZ KARACA, K., BRITO, D. V. C. & OLIVEIRA, A. M. M. 2019. MeCP2: A Critical Regulator of Chromatin in Neurodevelopment and Adult Brain Function. *Int J Mol Sci,* 20.

GUSSOW, A. B., PETROVSKI, S., WANG, Q., ALLEN, A. S. & GOLDSTEIN, D. B. 2016. The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. *Genome Biol,* 17**,** 9.

GUY, J., CHEVAL, H., SELFRIDGE, J. & BIRD, A. 2011. The role of MeCP2 in the brain. *Annu Rev Cell Dev Biol,* 27**,** 631-52.

GUY, J., HENDRICH, B., HOLMES, M., MARTIN, J. E. & BIRD, A. 2001. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat Genet,* 27**,** 322-6.

HAGBERG, B., AICARDI, J., DIAS, K. & RAMOS, O. 1983. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol,* 14**,** 471-9.

HAGBERG, B., HANEFELD, F., PERCY, A. & SKJELDAL, O. 2002. An update on clinically applicable diagnostic criteria in Rett syndrome. Comments to Rett Syndrome Clinical Criteria Consensus Panel Satellite to European Paediatric Neurology Society Meeting, Baden Baden, Germany, 11 September 2001. *Eur J Paediatr Neurol,* 6**,** 293-7.

HAGBERG, B. A. & SKJELDAL, O. H. 1994. Rett variants: a suggested model for inclusion criteria. *Pediatr Neurol,* 11**,** 5-11.

HAM, A. L., KUMAR, A., DEETER, R. & SCHANEN, N. C. 2005. Does genotype predict phenotype in Rett syndrome? *J Child Neurol,* 20**,** 768-78.

HANEFELD, F. 1985. The clinical pattern of the Rett syndrome. *Brain Dev,* 7**,** 320-5.

HARA, M., TAKAHASHI, T., MITSUMASU, C., IGATA, S., TAKANO, M., MINAMI, T., YASUKAWA, H., OKAYAMA, S., NAKAMURA, K., OKABE, Y., TANAKA, E., TAKEMURA, G., KOSAI, K., YAMASHITA, Y. & MATSUISHI, T. 2015. Disturbance of cardiac gene expression and cardiomyocyte structure predisposes Mecp2-null mice to arrhythmias. *Sci Rep,* 5**,** 11204.

HARDWICK, S. A., REUTER, K., WILLIAMSON, S. L., VASUDEVAN, V., DONALD, J., SLATER, K., BENNETTS, B., BEBBINGTON, A., LEONARD, H., WILLIAMS, S. R., SMITH, R. L., CLOOSTERMAN, D. & CHRISTODOULOU, J. 2007. Delineation of large deletions of the MECP2 gene in Rett syndrome patients, including a familial case with a male proband. *Eur J Hum Genet,* 15**,** 1218-29.

HECTOR, R. D., KALSCHEUER, V. M., HENNIG, F., LEONARD, H., DOWNS, J., CLARKE, A., BENKE, T. A., ARMSTRONG, J., PINEDA, M., BAILEY, M. E. S. & COBB, S. R. 2017. CDKL5 variants: Improving our understanding of a rare neurologic disorder. *Neurol Genet,* 3**,** e200.

HENDRICH, B. & TWEEDIE, S. 2003. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet,* 19**,** 269-77.

HOFFBUHR, K., DEVANEY, J. M., LAFLEUR, B., SIRIANNI, N., SCACHERI, C., GIRON, J., SCHUETTE, J., INNIS, J., MARINO, M., PHILIPPART, M., NARAYANAN, V., UMANSKY, R., KRONN, D., HOFFMAN, E. P. & NAIDU, S. 2001. MeCP2 mutations in children with and without the phenotype of Rett syndrome. *Neurology,* 56**,** 1486-95.

HOFFJAN, S., IBISLER, A., TSCHENTSCHER, A., DEKOMIEN, G., BIDINOST, C. & ROSA, A. L. 2016. WDR45 mutations in Rett (-like) syndrome and developmental delay: Case report and an appraisal of the literature. *Mol Cell Probes,* 30**,** 44-9.

HUPPKE, P. & GARTNER, J. 2005. Molecular diagnosis of Rett syndrome. *J Child Neurol,* 20**,** 732-6.

IP, J. P. K., MELLIOS, N. & SUR, M. 2018. Rett syndrome: insights into genetic, molecular and circuit mechanisms. *Nat Rev Neurosci,* 19**,** 368-382.

ISLAM, M. P., HERMAN, G. E. & DE LOS REYES, E. C. 2018. Language Regression in an Atypical SLC6A1 Mutation. *Semin Pediatr Neurol,* 26**,** 25-27.

ITOH, M., IDE, S., TAKASHIMA, S., KUDO, S., NOMURA, Y., SEGAWA, M., KUBOTA, T., MORI, H., TANAKA, S., HORIE, H., TANABE, Y. & GOTO, Y. 2007. Methyl CpG-binding protein 2 (a mutation of which causes Rett syndrome) directly regulates insulin-like growth factor binding protein 3 in mouse and human brains. *J Neuropathol Exp Neurol,* 66**,** 117-23.

IWAMA, K., MIZUGUCHI, T., TAKESHITA, E., NAKAGAWA, E., OKAZAKI, T., NOMURA, Y., IIJIMA, Y., KAJIURA, I., SUGAI, K., SAITO, T., SASAKI, M., YUGE, K., SAIKUSA, T., OKAMOTO, N., TAKAHASHI, S., AMAMOTO, M., TOMITA, I., KUMADA, S., ANZAI, Y., HOSHINO, K., FATTAL-VALEVSKI, A., SHIROMA, N., OHFU, M., MOROTO, M., TANDA, K., NAKAGAWA, T., SAKAKIBARA, T., NABATAME, S., MATSUO, M., YAMAMOTO, A., YUKISHITA, S., INOUE, K., WAGA, C., NAKAMURA, Y., WATANABE, S., OHBA, C., SENGOKU, T., FUJITA, A., MITSUHASHI, S., MIYATAKE, S., TAKATA, A., MIYAKE, N., OGATA, K., ITO, S., SAITSU, H., MATSUISHI, T., GOTO, Y. I. & MATSUMOTO, N. 2019. Genetic landscape of Rett syndrome-like phenotypes revealed by whole exome sequencing. *J Med Genet.*

JOHANNESEN, K. M., GARDELLA, E., LINNANKIVI, T., COURAGE, C., DE SAINT MARTIN, A., LEHESJOKI, A. E., MIGNOT, C., AFENJAR, A., LESCA, G., ABI-WARDE, M. T., CHELLY, J., PITON, A., MERRITT, J. L., 2ND, RODAN, L. H., TAN, W. H., BIRD, L. M., NESPECA, M., GLEESON, J. G., YOO, Y., CHOI, M., CHAE, J. H., CZAPANSKY-BEILMAN, D., REICHERT, S. C., PENDZIWIAT, M., VERHOEVEN, J. S., SCHELHAAS, H. J., DEVINSKY, O., CHRISTENSEN, J., SPECCHIO, N., TRIVISANO, M., WEBER, Y. G., NAVA, C., KEREN, B., DOUMMAR, D., SCHAEFER, E., HOPKINS, S., DUBBS, H., SHAW, J. E., PISANI, L., MYERS, C. T., TANG, S., TANG, S., PAL, D. K., MILLICHAP, J. J., CARVILL, G. L., HELBIG, K. L., MECARELLI, O., STRIANO, P., HELBIG, I., RUBBOLI, G., MEFFORD, H. C. & MOLLER, R. S. 2018. Defining the phenotypic spectrum of SLC6A1 mutations. *Epilepsia,* 59**,** 389-402.

JONES, P. L., VEENSTRA, G. J., WADE, P. A., VERMAAK, D., KASS, S. U., LANDSBERGER, N., STROUBOULIS, J. & WOLFFE, A. P. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet,* 19**,** 187-91.

KALSCHEUER, V. M., FREUDE, K., MUSANTE, L., JENSEN, L. R., YNTEMA, H. G., GECZ, J., SEFIANI, A., HOFFMANN, K., MOSER, B., HAAS, S., GUROK, U., HAESLER, S., ARANDA, B., NSHEDJAN, A., TZSCHACH, A., HARTMANN, N., ROLOFF, T. C., SHOICHET, S., HAGENS, O., TAO, J., VAN BOKHOVEN, H., TURNER, G., CHELLY, J., MORAINE, C., FRYNS, J. P., NUBER, U., HOELTZENBEIN, M., SCHARFF, C., SCHERTHAN, H., LENZNER, S., HAMEL, B. C., SCHWEIGER, S. &

ROPERS, H. H. 2003. Mutations in the polyglutamine binding protein 1 gene cause X-linked mental retardation. *Nat Genet,* 35**,** 313-5.

KATO, M., YAMAGATA, T., KUBOTA, M., ARAI, H., YAMASHITA, S., NAKAGAWA, T., FUJII, T., SUGAI, K., IMAI, K., USTER, T., CHITAYAT, D., WEISS, S., KASHII, H., KUSANO, R., MATSUMOTO, A., NAKAMURA, K., OYAZATO, Y., MAENO, M., NISHIYAMA, K., KODERA, H., NAKASHIMA, M., TSURUSAKI, Y., MIYAKE, N., SAITO, K., HAYASAKA, K., MATSUMOTO, N. & SAITSU, H. 2013. Clinical spectrum of early onset epileptic encephalopathies caused by KCNQ2 mutation. *Epilepsia,* 54**,** 1282-7.

KATZ, D. M., BIRD, A., COENRAADS, M., GRAY, S. J., MENON, D. U., PHILPOT, B. D. & TARQUINIO, D. C. 2016. Rett Syndrome: Crossing the Threshold to Clinical Translation. *Trends Neurosci,* 39**,** 100-113.

KENNEDY, B., KRONENBERG, Z., HU, H., MOORE, B., FLYGARE, S., REESE, M. G., JORDE, L. B., YANDELL, M. & HUFF, C. 2014. Using VAAST to Identify Disease-Associated Variants in Next-Generation Sequencing Data. *Curr Protoc Hum Genet,* 81**,** 6 14 1-25.

KERR, A. M. 1995. Early clinical signs in the Rett disorder. *Neuropediatrics,* 26**,** 67-71.

KERR, A. M., ARMSTRONG, D. D., PRESCOTT, R. J., DOYLE, D. & KEARNEY, D. L. 1997. Rett syndrome: analysis of deaths in the British survey. *Eur Child Adolesc Psychiatry,* 6 Suppl 1**,** 71-4.

KHWAJA, O. S., HO, E., BARNES, K. V., O'LEARY, H. M., PEREIRA, L. M., FINKELSTEIN, Y., NELSON, C. A., 3RD, VOGEL-FARLEY, V., DEGREGORIO, G., HOLM, I. A., KHATWA, U., KAPUR, K., ALEXANDER, M. E., FINNEGAN, D. M., CANTWELL, N. G., WALCO, A. C., RAPPAPORT, L., GREGAS, M., FICHOROVA, R. N., SHANNON, M. W., SUR, M. & KAUFMANN, W. E. 2014. Safety, pharmacokinetics, and preliminary assessment of efficacy of mecasermin (recombinant human IGF-1) for the treatment of Rett syndrome. *Proc Natl Acad Sci U S A,* 111**,** 4596-601.

KIM, K. Y., HYSOLLI, E. & PARK, I. H. 2011. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. *Proc Natl Acad Sci U S A,* 108**,** 14169-74.

KLAUCK, S. M., LINDSAY, S., BEYER, K. S., SPLITT, M., BURN, J. & POUSTKA, A. 2002. A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome. *Am J Hum Genet,* 70**,** 1034-7.

KLOSE, R. J. & BIRD, A. P. 2004. MeCP2 behaves as an elongated monomer that does not stably associate with the Sin3a chromatin remodeling complex. *J Biol Chem,* 279**,** 46490-6.

KNUDSEN, G. P., NEILSON, T. C., PEDERSEN, J., KERR, A., SCHWARTZ, M., HULTEN, M., BAILEY, M. E. & ORSTAVIK, K. H. 2006. Increased skewing of X chromosome inactivation in Rett syndrome patients and their mothers. *Eur J Hum Genet,* 14**,** 1189-94.

KOBOLDT, D. C., STEINBERG, K. M., LARSON, D. E., WILSON, R. K. & MARDIS, E. R. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell,* 155**,** 27-38.

KORTUM, F., DAS, S., FLINDT, M., MORRIS-ROSENDAHL, D. J., STEFANOVA, I., GOLDSTEIN, A., HORN, D., KLOPOCKI, E., KLUGER, G., MARTIN, P., RAUCH, A., ROUMER, A., SAITTA, S., WALSH, L. E., WIECZOREK, D., UYANIK, G., KUTSCHE, K. & DOBYNS, W. B. 2011. The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *J Med Genet,* 48**,** 396-406.

KRISHNARAJ, R., HO, G. & CHRISTODOULOU, J. 2017. RettBASE: Rett syndrome database update. *Hum Mutat,* 38**,** 922-931.

KRON, M., HOWELL, C. J., ADAMS, I. T., RANSBOTTOM, M., CHRISTIAN, D., OGIER, M. & KATZ, D. M. 2012. Brain activity mapping in Mecp2 mutant mice reveals functional deficits in forebrain circuits, including key nodes in the default mode network, that are reversed with ketamine treatment. *J Neurosci,* 32**,** 13860-72.

KYLE, S. M., VASHI, N. & JUSTICE, M. J. 2018. Rett syndrome: a neurological disorder with metabolic components. *Open Biol,* 8.

LADUCA, H., FARWELL, K. D., VUONG, H., LU, H. M., MU, W., SHAHMIRZADI, L., TANG, S., CHEN, J., BHIDE, S. & CHAO, E. C. 2017. Exome sequencing covers >98% of mutations identified on targeted next generation sequencing panels. *PLoS One,* 12**,** e0170843.

LEONARD, H. & BOWER, C. 1998. Is the girl with Rett syndrome normal at birth? *Dev Med Child Neurol,* 40**,** 115-21.

LEONARD, H., SILBERSTEIN, J., FALK, R., HOUWINK-MANVILLE, I., ELLAWAY, C., RAFFAELE, L. S., ENGERSTROM, I. W. & SCHANEN, C. 2001. Occurrence of Rett syndrome in boys. *J Child Neurol,* 16**,** 333-8.

LEWIS, J. D., MEEHAN, R. R., HENZEL, W. J., MAURER-FOGY, I., JEPPESEN, P., KLEIN, F. & BIRD, A. 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell,* 69**,** 905-14.

LIYANAGE, V. R. & RASTEGAR, M. 2014. Rett syndrome and MeCP2. *Neuromolecular Med,* 16**,** 231-64.

LOMBARDI, L. M., BAKER, S. A. & ZOGHBI, H. Y. 2015. MECP2 disorders: from the clinic to mice and back. *J Clin Invest,* 125**,** 2914-23.

LOPES, F., BARBOSA, M., AMEUR, A., SOARES, G., DE SA, J., DIAS, A. I., OLIVEIRA, G., CABRAL, P., TEMUDO, T., CALADO, E., CRUZ, I. F., VIEIRA, J. P., OLIVEIRA, R., ESTEVES, S., SAUER, S., JONASSON, I., SYVANEN, A. C., GYLLENSTEN, U., PINTO, D. & MACIEL, P. 2016. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet,* 53**,** 190-9.

LUPSKI, J. R., REID, J. G., GONZAGA-JAUREGUI, C., RIO DEIROS, D., CHEN, D. C., NAZARETH, L., BAINBRIDGE, M., DINH, H., JING, C., WHEELER, D. A., MCGUIRE, A. L., ZHANG, F., STANKIEWICZ, P., HALPERIN, J. J., YANG, C., GEHMAN, C., GUO, D., IRIKAT, R. K., TOM, W., FANTIN, N. J., MUZNY, D. M. & GIBBS, R. A. 2010. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med,* 362**,** 1181-91.

LYST, M. J., EKIERT, R., EBERT, D. H., MERUSI, C., NOWAK, J., SELFRIDGE, J., GUY, J., KASTAN, N. R., ROBINSON, N. D., DE LIMA ALVES, F., RAPPSILBER, J., GREENBERG, M. E. & BIRD, A. 2013. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat Neurosci,* 16**,** 898-902.

MADSEN, K. K., HANSEN, G. H., DANIELSEN, E. M. & SCHOUSBOE, A. 2015. The subcellular localization of GABA transporters and its implication for seizure management. *Neurochem Res,* 40**,** 410-9.

MARI, F., AZIMONTI, S., BERTANI, I., BOLOGNESE, F., COLOMBO, E., CASELLI, R., SCALA, E., LONGO, I., GROSSO, S., PESCUCCI, C., ARIANI, F., HAYEK, G., BALESTRI, P., BERGO, A., BADARACCO, G., ZAPPELLA, M., BROCCOLI, V., RENIERI, A., KILSTRUP-NIELSEN, C. & LANDSBERGER, N. 2005. CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Hum Mol Genet,* 14**,** 1935-46.

MARSCHIK, P. B., LEMCKE, S., EINSPIELER, C., ZHANG, D., BOLTE, S., TOWNEND, G. S. & LAURITSEN, M. B. 2018. Early development in Rett syndrome - the benefits and difficulties of a birth cohort approach. *Dev Neurorehabil,* 21**,** 68-72.

MARSHALL, B., ISIDRO, G. & BOAVIDA, M. G. 1996. Insertion of a short Alu sequence into the hMSH2 gene following a double cross over next to sequences with chi homology. *Gene,* 174**,** 175-9.

MARX, V. 2013. Next-generation sequencing: The genome jigsaw. *Nature,* 501**,** 263-8.

MATTISON, K. A., BUTLER, K. M., INGLIS, G. A. S., DAYAN, O., BOUSSIDAN, H., BHAMBHANI, V., PHILBROOK, B., DA SILVA, C., ALEXANDER, J. J., KANNER, B. I. & ESCAYG, A. 2018. SLC6A1 variants identified in epilepsy patients reduce gamma-aminobutyric acid transport. *Epilepsia,* 59**,** e135-e141.

MCCAULEY, M. D., WANG, T., MIKE, E., HERRERA, J., BEAVERS, D. L., HUANG, T. W., WARD, C. S., SKINNER, S., PERCY, A. K., GLAZE, D. G., WEHRENS, X. H. & NEUL, J. L. 2011. Pathogenesis of lethal cardiac arrhythmias in Mecp2 mutant mice: implication for therapy in Rett syndrome. *Sci Transl Med,* 3**,** 113ra125.

MEEHAN, R. R., LEWIS, J. D. & BIRD, A. P. 1992. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res,* 20**,** 5085-92.

MELLEN, M., AYATA, P., DEWELL, S., KRIAUCIONIS, S. & HEINTZ, N. 2012. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell,* 151**,** 1417-30.

MELLIOS, N., FELDMAN, D. A., SHERIDAN, S. D., IP, J. P. K., KWOK, S., AMOAH, S. K., ROSEN, B., RODRIGUEZ, B. A., CRAWFORD, B., SWAMINATHAN, R., CHOU, S., LI, Y., ZIATS, M., ERNST, C., JAENISCH, R., HAGGARTY, S. J. & SUR, M. 2018. MeCP2-regulated miRNAs control early human neurogenesis through differential effects on ERK and AKT signaling. *Mol Psychiatry,* 23**,** 1051-1065.

MEYERSON, M., GABRIEL, S. & GETZ, G. 2010. Advances in understanding cancer genomes through second-generation sequencing. *Nat Rev Genet,* 11**,** 685-96.

MIRZAA, G. M., PACIORKOWSKI, A. R., MARSH, E. D., BERRY-KRAVIS, E. M., MEDNE, L., ALKHATEEB, A., GRIX, A., WIRRELL, E. C., POWELL, B. R., NICKELS, K. C., BURTON, B., PARAS, A., KIM, K., CHUNG, W., DOBYNS, W. B. & DAS, S. 2013. CDKL5 and ARX mutations in males with early-onset epilepsy. *Pediatr Neurol,* 48**,** 367-77.

MITTER, D., PRINGSHEIM, M., KAULISCH, M., PLUMACHER, K. S., SCHRODER, S., WARTHEMANN, R., ABOU JAMRA, R., BAETHMANN, M., BAST, T., BUTTEL, H. M., COHEN, J. S., CONOVER, E., COURAGE, C., EGER, A., FATEMI, A., GREBE, T. A., HAUSER, N. S., HEINRITZ, W., HELBIG, K. L., HERUTH, M., HUHLE, D., HOFT, K., KARCH, S., KLUGER, G., KORENKE, G. C., LEMKE, J. R., LUTZ, R. E., PATZER, S., PREHL, I., HOERTNAGEL, K., RAMSEY, K., RATING, T., RIESS, A., ROHENA, L., SCHIMMEL, M., WESTMAN, R., ZECH, F. M., ZOLL, B., MALZAHN, D., ZIRN, B. & BROCKMANN, K. 2018. FOXG1 syndrome: genotype-phenotype association in 83 patients with FOXG1 variants. *Genet Med,* 20**,** 98-108.

MONROS, E., ARMSTRONG, J., AIBAR, E., POO, P., CANOS, I. & PINEDA, M. 2001. Rett syndrome in Spain: mutation analysis and clinical correlations. *Brain Dev,* 23 Suppl 1**,** S251-3.

MOUNT, R. H., CHARMAN, T., HASTINGS, R. P., REILLY, S. & CASS, H. 2002. The Rett Syndrome Behaviour Questionnaire (RSBQ): refining the behavioural phenotype of Rett syndrome. *J Child Psychol Psychiatry,* 43**,** 1099-110.

NACHMAN, M. W. & CROWELL, S. L. 2000. Estimate of the mutation rate per nucleotide in humans. *Genetics,* 156**,** 297-304.

NAIDU, S. 1997. Rett syndrome: a disorder affecting early brain growth. *Ann Neurol,* 42**,** 3-10.

NAN, X. & BIRD, A. 2001. The biological functions of the methyl-CpG-binding protein MeCP2 and its implication in Rett syndrome. *Brain Dev,* 23 Suppl 1**,** S32-7.

NEUL, J. L., FANG, P., BARRISH, J., LANE, J., CAEG, E. B., SMITH, E. O., ZOGHBI, H., PERCY, A. & GLAZE, D. G. 2008. Specific mutations in methyl-CpG-binding protein 2 confer different severity in Rett syndrome. *Neurology,* 70**,** 1313-21.

NEUL, J. L., KAUFMANN, W. E., GLAZE, D. G., CHRISTODOULOU, J., CLARKE, A. J., BAHI-BUISSON, N., LEONARD, H., BAILEY, M. E., SCHANEN, N. C., ZAPPELLA, M., RENIERI, A., HUPPKE, P., PERCY, A. K. & RETTSEARCH, C. 2010. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol,* 68**,** 944-50.

NOMURA, Y. & SEGAWA, M. 1992. Motor symptoms of the Rett syndrome: abnormal muscle tone, posture, locomotion and stereotyped movement. *Brain Dev,* 14 Suppl**,** S21-8.

O'LEARY, H. M., KAUFMANN, W. E., BARNES, K. V., RAKESH, K., KAPUR, K., TARQUINIO, D. C., CANTWELL, N. G., ROCHE, K. J., ROSE, S. A., WALCO, A. C., BRUCK, N. M., BAZIN, G. A., HOLM, I. A., ALEXANDER, M. E., SWANSON, L. C., BACZEWSKI, L. M., POON, C., MAYOR TORRES, J. M., NELSON, C. A., 3RD & SAHIN, M. 2018. Placebo-controlled crossover assessment of mecasermin for the treatment of Rett syndrome. *Ann Clin Transl Neurol,* 5**,** 323-332.

OGIER, M. & KATZ, D. M. 2008. Breathing dysfunction in Rett syndrome: understanding epigenetic regulation of the respiratory network. *Respir Physiol Neurobiol,* 164**,** 55-63.

OLSON, C. O., ZACHARIAH, R. M., EZEONWUKA, C. D., LIYANAGE, V. R. & RASTEGAR, M. 2014. Brain region-specific expression of MeCP2 isoforms correlates with DNA methylation within Mecp2 regulatory elements. *PLoS One,* 9**,** e90645.

OLSON, H. E., TAMBUNAN, D., LACOURSIERE, C., GOLDENBERG, M., PINSKY, R., MARTIN, E., HO, E., KHWAJA, O., KAUFMANN, W. E. & PODURI, A. 2015. Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. *Am J Med Genet A,* 167A**,** 2017-25.

OREFICE, L. L., ZIMMERMAN, A. L., CHIRILA, A. M., SLEBODA, S. J., HEAD, J. P. & GINTY, D. D. 2016. Peripheral Mechanosensory Neuron Dysfunction Underlies Tactile and Behavioral Deficits in Mouse Models of ASDs. *Cell,* 166**,** 299-313.

ORRICO, A., LAM, C., GALLI, L., DOTTI, M. T., HAYEK, G., TONG, S. F., POON, P. M., ZAPPELLA, M., FEDERICO, A. & SORRENTINO, V. 2000. MECP2 mutation in male patients with non-specific X-linked mental retardation. *FEBS Lett,* 481**,** 285-8.

PETAZZI, P., SANDOVAL, J., SZCZESNA, K., JORGE, O. C., ROA, L., SAYOLS, S., GOMEZ, A., HUERTAS, D. & ESTELLER, M. 2013. Dysregulation of the long non-coding RNA transcriptome in a Rett syndrome mouse model. *RNA Biol,* 10**,** 1197-203.

PETROVSKI, S., WANG, Q., HEINZEN, E. L., ALLEN, A. S. & GOLDSTEIN, D. B. 2013. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet,* 9**,** e1003709.

POKORNY, F. B., BARTL-POKORNY, K. D., EINSPIELER, C., ZHANG, D., VOLLMANN, R., BOLTE, S., GUGATSCHKA, M., SCHULLER, B. W. & MARSCHIK, P. B. 2018. Typical vs. atypical: Combining auditory Gestalt perception and acoustic analysis of early vocalisations in Rett syndrome. *Res Dev Disabil,* 82**,** 109-119.

RAMIREZ, J. M., WARD, C. S. & NEUL, J. L. 2013. Breathing challenges in Rett syndrome: lessons learned from humans and animal models. *Respir Physiol Neurobiol,* 189**,** 280-7.

RAUCH, A., WIECZOREK, D., GRAF, E., WIELAND, T., ENDELE, S., SCHWARZMAYR, T., ALBRECHT, B., BARTHOLDI, D., BEYGO, J., DI DONATO, N., DUFKE, A., CREMER, K., HEMPEL, M., HORN, D., HOYER, J., JOSET, P., ROPKE, A., MOOG, U., RIESS, A., THIEL, C. T., TZSCHACH, A., WIESENER, A., WOHLLEBER, E., ZWEIER, C., EKICI, A. B., ZINK, A. M., RUMP, A., MEISINGER, C., GRALLERT, H., STICHT, H., SCHENCK, A., ENGELS, H., RAPPOLD, G., SCHROCK, E., WIEACKER, P., RIESS, O., MEITINGER, T., REIS, A. & STROM, T. M. 2012. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet,* 380**,** 1674-82.

REHM, H. L., BALE, S. J., BAYRAK-TOYDEMIR, P., BERG, J. S., BROWN, K. K., DEIGNAN, J. L., FRIEZ, M. J., FUNKE, B. H., HEGDE, M. R., LYON, E., WORKING GROUP OF THE AMERICAN COLLEGE OF MEDICAL, G. & GENOMICS LABORATORY QUALITY ASSURANCE, C. 2013. ACMG clinical laboratory standards for next-generation sequencing. *Genet Med,* 15**,** 733-47.

RENTHAL, W., BOXER, L. D., HRVATIN, S., LI, E., SILBERFELD, A., NAGY, M. A., GRIFFITH, E. C., VIERBUCHEN, T. & GREENBERG, M. E. 2018. Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nat Neurosci,* 21**,** 1670-1679.

RETT, A. 1966. [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien Med Wochenschr,* 116**,** 723-6.

RICCERI, L., DE FILIPPIS, B. & LAVIOLA, G. 2008. Mouse models of Rett syndrome: from behavioural phenotyping to preclinical evaluation of new therapeutic approaches. *Behav Pharmacol,* 19**,** 501-17.

RICHARDS, S., AZIZ, N., BALE, S., BICK, D., DAS, S., GASTIER-FOSTER, J., GRODY, W. W., HEGDE, M., LYON, E., SPECTOR, E., VOELKERDING, K., REHM, H. L. & COMMITTEE, A. L. Q. A. 2015. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med,* 17**,** 405-24.

ROCA, I., FERNANDEZ-MARMIESSE, A., GOUVEIA, S., SEGOVIA, M. & COUCE, M. L. 2018. Prioritization of Variants Detected by Next Generation Sequencing According to the Mutation Tolerance and Mutational Architecture of the Corresponding Genes. *Int J Mol Sci,* 19.

ROLANDO, S. 1985. Rett syndrome: report of eight cases. *Brain Dev,* 7**,** 290-6.

ROSS, P. D., GUY, J., SELFRIDGE, J., KAMAL, B., BAHEY, N., TANNER, K. E., GILLINGWATER, T. H., JONES, R. A., LOUGHREY, C. M., MCCARROLL, C. S., BAILEY, M. E., BIRD, A. & COBB, S. 2016. Exclusive expression of MeCP2 in the nervous system distinguishes between brain and peripheral Rett syndrome-like phenotypes. *Hum Mol Genet,* 25**,** 4389-4404.

RUSSO, S., MARCHI, M., COGLIATI, F., BONATI, M. T., PINTAUDI, M., VENESELLI, E., SALETTI, V., BALESTRINI, M., BEN-ZEEV, B. & LARIZZA, L. 2009. Novel mutations in the CDKL5 gene, predicted effects and associated phenotypes. *Neurogenetics,* 10**,** 241-50.

SAITSU, H., KATO, M., SHIMONO, M., SENJU, A., TANABE, S., KIMURA, T., NISHIYAMA, K., YONEDA, Y., KONDO, Y., TSURUSAKI, Y., DOI, H., MIYAKE, N., HAYASAKA, K. & MATSUMOTO, N. 2012. Association of genomic deletions in the STXBP1 gene with Ohtahara syndrome. *Clin Genet,* 81**,** 399-402.

SAITSU, H., TOHYAMA, J., KUMADA, T., EGAWA, K., HAMADA, K., OKADA, I., MIZUGUCHI, T., OSAKA, H., MIYATA, R., FURUKAWA, T., HAGINOYA, K., HOSHINO, H., GOTO, T., HACHIYA, Y., YAMAGATA, T., SAITOH, S., NAGAI, T., NISHIYAMA, K., NISHIMURA, A., MIYAKE, N., KOMADA, M., HAYASHI, K., HIRAI, S., OGATA, K., KATO, M., FUKUDA, A. & MATSUMOTO, N. 2010. Dominant-negative mutations in alpha-II spectrin cause West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay. *Am J Hum Genet,* 86**,** 881-91.

SAJAN, S. A., JHANGIANI, S. N., MUZNY, D. M., GIBBS, R. A., LUPSKI, J. R., GLAZE, D. G., KAUFMANN, W. E., SKINNER, S. A., ANNESE, F., FRIEZ, M. J., LANE, J., PERCY, A. K. & NEUL, J. L. 2017. Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet Med,* 19**,** 13-19.

SAMACO, R. C., MCGRAW, C. M., WARD, C. S., SUN, Y., NEUL, J. L. & ZOGHBI, H. Y. 2013. Female Mecp2(+/-) mice display robust behavioral deficits on two different genetic backgrounds providing a framework for pre-clinical studies. *Hum Mol Genet,* 22**,** 96-109.

SAMANTA, D. & ZARATE, Y. A. 2019. Widening phenotypic spectrum of GABBR2 mutation. *Acta Neurol Belg,* 119**,** 493-496.

SANDERS, S. J., CAMPBELL, A. J., COTTRELL, J. R., MOLLER, R. S., WAGNER, F. F., AULDRIDGE, A. L., BERNIER, R. A., CATTERALL, W. A., CHUNG, W. K., EMPFIELD, J. R., GEORGE, A. L., JR., HIPP, J. F., KHWAJA, O., KISKINIS, E., LAL, D., MALHOTRA, D., MILLICHAP, J. J., OTIS, T. S., PETROU, S., PITT, G., SCHUST, L. F., TAYLOR, C. M., TJERNAGEL, J., SPIRO, J. E. & BENDER, K. J. 2018. Progress in Understanding and Treating SCN2A-Mediated Disorders. *Trends Neurosci,* 41**,** 442-456.

SCHONEWOLF-GREULICH, B., BISGAARD, A. M., MOLLER, R. S., DUNO, M., BRONDUM-NIELSEN, K., KAUR, S., VAN BERGEN, N. J., LUNKE, S., EGGERS, S., JESPERSGAARD, C., CHRISTODOULOU, J. & TUMER, Z. 2019. Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin Genet,* 95**,** 221-230.

SHOVLIN, S. & TROPEA, D. 2018. Transcriptome level analysis in Rett syndrome using human samples from different tissues. *Orphanet J Rare Dis,* 13**,** 113.

SHULYAKOVA, N., ANDREAZZA, A. C., MILLS, L. R. & EUBANKS, J. H. 2017. Mitochondrial Dysfunction in the Pathogenesis of Rett Syndrome: Implications for Mitochondria-Targeted Therapies. *Front Cell Neurosci,* 11**,** 58.

SMITH, N. G., WEBSTER, M. T. & ELLEGREN, H. 2002. Deterministic mutation rate variation in the human genome. *Genome Res,* 12**,** 1350-6.

SRIVASTAVA, S., DESAI, S., COHEN, J., SMITH-HICKS, C., BARANANO, K., FATEMI, A. & NAIDU, S. 2018. Monogenic disorders that mimic the phenotype of Rett syndrome. *Neurogenetics,* 19**,** 41-47.

SUTER, B., TREADWELL-DEERING, D., ZOGHBI, H. Y., GLAZE, D. G. & NEUL, J. L. 2014. Brief report: MECP2 mutations in people without Rett syndrome. *J Autism Dev Disord,* 44**,** 703-11.

TAO, J., VAN ESCH, H., HAGEDORN-GREIWE, M., HOFFMANN, K., MOSER, B., RAYNAUD, M., SPERNER, J., FRYNS, J. P., SCHWINGER, E., GECZ, J., ROPERS, H. H. & KALSCHEUER, V. M. 2004. Mutations in the X-linked cyclin-dependent kinase-like 5 (CDKL5/STK9) gene are associated with severe neurodevelopmental retardation. *Am J Hum Genet,* 75**,** 1149-54.

TARQUINIO, D. C., HOU, W., NEUL, J. L., KAUFMANN, W. E., GLAZE, D. G., MOTIL, K. J., SKINNER, S. A., LEE, H. S. & PERCY, A. K. 2015. The Changing Face of Survival in Rett Syndrome and MECP2-Related Disorders. *Pediatr Neurol,* 53**,** 402-11.

TATE, P., SKARNES, W. & BIRD, A. 1996. The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. *Nat Genet,* 12**,** 205-8.

TEJADA, M. I., PENAGARIKANO, O., RODRIGUEZ-REVENGA, L., MARTINEZ-BOUZAS, C., GARCIA, B., BADENAS, C., GUITART, M., MINGUEZ, M., GARCIA-ALEGRIA, E., SANZ-PARRA, A., BERISTAIN, E. & MILA, M. 2006. Screening for MECP2 mutations in Spanish patients with an unexplained mental retardation. *Clin Genet,* 70**,** 140-4.

TEMUDO, T., SANTOS, M., RAMOS, E., DIAS, K., VIEIRA, J. P., MOREIRA, A., CALADO, E., CARRILHO, I., OLIVEIRA, G., LEVY, A., BARBOT, C., FONSECA, M., CABRAL, A., CABRAL, P., MONTEIRO, J., BORGES, L., GOMES, R., MIRA, G., PEREIRA, S. A., SANTOS, M., FERNANDES, A., EPPLEN, J. T., SEQUEIROS, J. & MACIEL, P. 2011. Rett syndrome with and without detected MECP2 mutations: an attempt to redefine phenotypes. *Brain Dev,* 33**,** 69-76.

TROPEA, D., GIACOMETTI, E., WILSON, N. R., BEARD, C., MCCURRY, C., FU, D. D., FLANNERY, R., JAENISCH, R. & SUR, M. 2009. Partial reversal of Rett Syndrome-like symptoms in MeCP2 mutant mice. *Proc Natl Acad Sci U S A,* 106**,** 2029-34.

URE, K., LU, H., WANG, W., ITO-ISHIDA, A., WU, Z., HE, L. J., SZTAINBERG, Y., CHEN, W., TANG, J. & ZOGHBI, H. Y. 2016. Restoration of Mecp2 expression in GABAergic neurons is sufficient to rescue multiple disease features in a mouse model of Rett syndrome. *Elife,* 5.

VAN DER AA, N., VAN DEN BERGH, M., PONOMARENKO, N., VERSTRAETE, L., CEULEMANS, B. & STORM, K. 2011. Analysis of FOXG1 Is Highly Recommended in Male and Female Patients with Rett Syndrome. *Mol Syndromol,* 1**,** 290-293.

VAN ESCH, H., BAUTERS, M., IGNATIUS, J., JANSEN, M., RAYNAUD, M., HOLLANDERS, K., LUGTENBERG, D., BIENVENU, T., JENSEN, L. R., GECZ, J., MORAINE, C., MARYNEN, P., FRYNS, J. P. & FROYEN, G. 2005. Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet,* 77**,** 442-53.

VASHI, N. & JUSTICE, M. J. 2019. Treating Rett syndrome: from mouse models to human therapies. *Mamm Genome,* 30**,** 90-110.

VEERARAGAVAN, S., WAN, Y. W., CONNOLLY, D. R., HAMILTON, S. M., WARD, C. S., SORIANO, S., PITCHER, M. R., MCGRAW, C. M., HUANG, S. G., GREEN, J. R., YUVA, L. A., LIANG, A. J., NEUL, J. L., YASUI, D. H., LASALLE, J. M., LIU, Z., PAYLOR, R. & SAMACO, R. C. 2016. Loss of MeCP2 in the rat models regression, impaired sociability and transcriptional deficits of Rett syndrome. *Hum Mol Genet,* 25**,** 3284-3302.

VEGAS, N., CAVALLIN, M., MAILLARD, C., BODDAERT, N., TOULOUSE, J., SCHAEFER, E., LERMAN-SAGIE, T., LEV, D., MAGALIE, B., MOUTTON, S., HAAN, E., ISIDOR, B., HERON, D., MILH, M., RONDEAU, S., MICHOT, C., VALENCE, S., WAGNER, S., HULLY, M., MIGNOT, C., MASUREL, A., DATTA, A., ODENT, S., NIZON, M., LAZARO, L., VINCENT, M., COGNE, B., GUERROT, A. M., ARPIN, S., PEDESPAN, J. M., CAUBEL, I., PONTIER, B., TROUDE, B., RIVIER, F., PHILIPPE, C., BIENVENU, T., SPITZ, M. A., BERY, A. & BAHI-BUISSON, N. 2018. Delineating FOXG1 syndrome: From congenital microcephaly to hyperkinetic encephalopathy. *Neurol Genet,* 4**,** e281.

VELTMAN, J. A. & BRUNNER, H. G. 2012. De novo mutations in human genetic disease. *Nat Rev Genet,* 13**,** 565-75.

VUILLAUME, M. L., JEANNE, M., XUE, L., BLESSON, S., DENOMME-PICHON, A. S., ALIROL, S., BRULARD, C., COLIN, E., ISIDOR, B., GILBERT-DUSSARDIER, B., ODENT, S., PARENT, P., DONNART, A., REDON, R., BEZIEAU, S., RONDARD, P., LAUMONNIER, F. & TOUTAIN, A. 2018. A novel mutation in the transmembrane 6 domain of GABBR2 leads to a Rett-like phenotype. *Ann Neurol,* 83**,** 437-439.

WANG, I. T., ALLEN, M., GOFFIN, D., ZHU, X., FAIRLESS, A. H., BRODKIN, E. S., SIEGEL, S. J., MARSH, E. D., BLENDY, J. A. & ZHOU, Z. 2012. Loss of CDKL5 disrupts kinome profile and event-related potentials leading to autistic-like phenotypes in mice. *Proc Natl Acad Sci U S A,* 109**,** 21516-21.

WANG, J., ZHANG, Q., CHEN, Y., YU, S., WU, X., BAO, X. & WEN, Y. 2018. Novel MEF2C point mutations in Chinese patients with Rett (-like) syndrome or non-syndromic intellectual disability: insights into genotype-phenotype correlation. *BMC Med Genet,* 19**,** 191.

WATSON, P., BLACK, G., RAMSDEN, S., BARROW, M., SUPER, M., KERR, B. & CLAYTON-SMITH, J. 2001. Angelman syndrome phenotype associated with mutations in MECP2, a gene encoding a methyl CpG binding protein. *J Med Genet,* 38**,** 224-8.

WEAVING, L. S., CHRISTODOULOU, J., WILLIAMSON, S. L., FRIEND, K. L., MCKENZIE, O. L., ARCHER, H., EVANS, J., CLARKE, A., PELKA, G. J., TAM, P. P., WATSON, C., LAHOOTI, H., ELLAWAY, C. J., BENNETTS, B., LEONARD, H. & GECZ, J. 2004. Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet,* 75**,** 1079-93.

WEAVING, L. S., ELLAWAY, C. J., GECZ, J. & CHRISTODOULOU, J. 2005. Rett syndrome: clinical review and genetic update. *J Med Genet,* 42**,** 1-7.

WECKHUYSEN, S., IVANOVIC, V., HENDRICKX, R., VAN COSTER, R., HJALGRIM, H., MOLLER, R. S., GRONBORG, S., SCHOONJANS, A. S., CEULEMANS, B., HEAVIN, S. B., ELTZE, C., HORVATH, R., CASARA, G., PISANO, T., GIORDANO, L., ROSTASY, K., HABERLANDT, E., ALBRECHT, B., BEVOT, A., BENKEL, I., SYRBE, S., SHEIDLEY, B., GUERRINI, R., PODURI, A., LEMKE, J. R., MANDELSTAM, S., SCHEFFER, I., ANGRIMAN, M., STRIANO, P., MARINI, C., SULS, A., DE JONGHE, P. & GROUP, K. S. 2013. Extending the KCNQ2 encephalopathy spectrum: clinical and neuroimaging findings in 17 patients. *Neurology,* 81**,** 1697-703.

WITT-ENGERSTROM, I. & GILLBERG, C. 1987. Rett syndrome in Sweden. *J Autism Dev Disord,* 17**,** 149-50.

WU, H., WANG, C., GREGORY, K. J., HAN, G. W., CHO, H. P., XIA, Y., NISWENDER, C. M., KATRITCH, V., MEILER, J., CHEREZOV, V., CONN, P. J. & STEVENS, R. C. 2014. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science,* 344**,** 58-64.

YASUI, D. H., GONZALES, M. L., AFLATOONI, J. O., CRARY, F. K., HU, D. J., GAVINO, B. J., GOLUB, M. S., VINCENT, J. B., CAROLYN SCHANEN, N., OLSON, C. O., RASTEGAR, M. & LASALLE, J. M. 2014. Mice with an isoform-ablating Mecp2 exon 1 mutation recapitulate the neurologic deficits of Rett syndrome. *Hum Mol Genet,* 23**,** 2447-58.

YASUI, D. H., PEDDADA, S., BIEDA, M. C., VALLERO, R. O., HOGART, A., NAGARAJAN, R. P., THATCHER, K. N., FARNHAM, P. J. & LASALLE, J. M. 2007. Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc Natl Acad Sci U S A,* 104**,** 19416-21.

YOO, Y., JUNG, J., LEE, Y. N., LEE, Y., CHO, H., NA, E., HONG, J., KIM, E., LEE, J. S., LEE, J. S., HONG, C., PARK, S. Y., WIE, J., MILLER, K., SHUR, N., CLOW, C., EBEL, R. S.,

DEBROSSE, S. D., HENDERSON, L. B., WILLAERT, R., CASTALDI, C., TIKHONOVA, I., BILGUVAR, K., MANE, S., KIM, K. J., HWANG, Y. S., LEE, S. G., SO, I., LIM, B. C., CHOI, H. J., SEONG, J. Y., SHIN, Y. B., JUNG, H., CHAE, J. H. & CHOI, M. 2017. GABBR2 mutations determine phenotype in rett syndrome and epileptic encephalopathy. *Ann Neurol,* 82**,** 466-478.

ZACHARIAH, R. M. & RASTEGAR, M. 2012. Linking epigenetics to human disease and Rett syndrome: the emerging novel and challenging concepts in MeCP2 research. *Neural Plast,* 2012**,** 415825.

ZAPPELLA, M. 1992. The Rett girls with preserved speech. *Brain Dev,* 14**,** 98-101.

ZEEV, B. B., YARON, Y., SCHANEN, N. C., WOLF, H., BRANDT, N., GINOT, N., SHOMRAT, R. & ORR-URTREGER, A. 2002. Rett syndrome: clinical manifestations in males with MECP2 mutations. *J Child Neurol,* 17**,** 20-4.

ZHANG, Q., YANG, X., WANG, J., LI, J., WU, Q., WEN, Y., ZHAO, Y., ZHANG, X., YAO, H., WU, X., YU, S., WEI, L. & BAO, X. 2019. Correction: Genomic mosaicism in the pathogenesis and inheritance of a Rett syndrome cohort. *Genet Med.*

ZHAO, M., WANG, Q., WANG, Q., JIA, P. & ZHAO, Z. 2013. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics,* 14 Suppl 11**,** S1.

ZHAO, N., MA, D., LEONG, W. Y., HAN, J., VANDONGEN, A., CHEN, T. & GOH, E. L. 2015. The methyl-CpG-binding domain (MBD) is crucial for MeCP2's dysfunction-induced defects in adult newborn neurons. *Front Cell Neurosci,* 9**,** 158.

ZLATANOVA, J. 2005. MeCP2: the chromatin connection and beyond. *Biochem Cell Biol,* 83**,** 251-62.

ZOGHBI, H. Y. 2005. MeCP2 dysfunction in humans and mice. *J Child Neurol,* 20**,** 736-40.

ZOLL, B., HUPPKE, P., WESSEL, A., BARTELS, I. & LACCONE, F. 2004. Fetal alcohol syndrome in association with Rett syndrome. *Genet Couns,* 15**,** 207-12.

ZWEIER, M., GREGOR, A., ZWEIER, C., ENGELS, H., STICHT, H., WOHLLEBER, E., BIJLSMA, E. K., HOLDER, S. E., ZENKER, M., ROSSIER, E., GRASSHOFF, U., JOHNSON, D. S., ROBERTSON, L., FIRTH, H. V., CORNELIA, K., EKICI, A. B., REIS, A. & RAUCH, A. 2010. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. *Hum Mutat,* 31**,** 722-33.

ANNEXO

Tabla complementaria

Tabla 1. Tratamientos preclínicos dirigidos a vías *downstream* de Mecp2. *Obtenida de Vashi N and Justice MJ 2019*

Treatment	Mechanism	Mouse model	Prolong lifespan	Improved phenotype	References	Clinical trial
Neurotransmitter signaling						
Citalopram	Serotonin reuptake blocker	Mecp2*tm1.1Bird*	NT	Improved sensitivity to carbon dioxide exposure	Toward et al. (2013)	–
8-OH-DPAT	Serotonin 1a agonist	Mecp2*tm1.1Bird*	NT	Reduced apneas	Abdala et al. (2010)	–
F15599	Serotonin 1a agonist	Mecp2*tm1.1Bird*	NT	Reduced apneas and improves breathing irregularity	Levitt et al. (2013)	–
Sarizotan	Serotonin 1a agonist & dopamine D2-like receptor	Mecp2*tm1.1Bird* Mecp2*tm1.1Jae* Mecp2*tm1.1Jtc*	NT	Reduced apneas and improves breathing irregularity	Abdala et al. (2014)	Y
LP-211	Serotonin 7 receptor agonist	Mecp2*tm1Hzo*	NT	Improved overall health, memory and anxiety	De Filippis et al. (2014)	–
Levodopa	Dopaminergic stimulation	Mecp2*tm1.1Bird*	Y	Improved motor activity	Szczesna et al. (2014)	–
Ketamine	NMDA receptor antagonist	Mecp2*tm1.1Jae*	NT	Improved startle response	Kron et al. (2012)	Y
Ketamine	NMDA receptor antagonist	Mecp2*tm1.1Bird*	Y	Improved limb clasping, motor coordination and reduced apneas	Patrizi et al. (2016)	Y
NO-711	GABA reuptake blocker	Mecp2*tm1.1Bird*	NT	Reduced apneas	Abdala et al. (2010)	–
Benzodiazepine diazepam	GABA reuptake blocker	Mecp2*tm1.1Bird*	NT	Reduced apneas	Abdala et al. (2010)	–
L-838,417	GABA reuptake blocker	Mecp2*tm1.1Bird*	NT	Reduced apneas	Abdala et al. (2010)	–
Tiagabine	GABA reuptake blocker	Mecp2*tm1.1Bird*	Y	No improvement	El-Khoury et al. (2014)	–

THIP	GABA reuptake blocker	Mecp2*tm1.1Bird*	Y	Improved motor function, social behavior, and reduced apneas	Zhong et al. (2016)	–
Mirtazapine	GABA reuptake blocker	Mecp2*tm1.1Bird*	NT	Improved overall healthy, neuronal morphology, dendritic spine number, anxiety	Bittolo et al. (2016)	–
VUO462807	mGlu5 positive allosteric modulator	Mecp2*tm1.1Bird*	N	Improved motor function and limb clasping	Gogliotti et al. (2016)	–
VU0422288	mGlu7 positive allosteric modulator	Mecp2*tm1.1Bird*	NT	Reduced apneas, improves learning and memory	Gogliotti et al. (2017)	–
CTEP	mGluR5 negative allosteric modulator	Mecp2*tm1.1Bird*	Y	Reduced apneas, improved memory	Tao et al. (2016)	–
Acetyl-L-carnitine	Acetyl group donor	Mecp2*tm1.1Jae*	NT	Improved weight gain, motor activity and memory	Schaevitz et al. (2012)	–
Acetyl-L-carnitine	Acetyl group donor	Mecp2*tm1.1Jae*	NT	Improved weight gain, motor activity and memory	Schaevitz et al. (2012)	–
Choline	ACh	Mecp2*tm1.1Jae*	NT	Improved motor coordination and activity	Nag and BergerSweeney (2007)	–
Choline	ACh	Mecp2*tm1Hzo*	NT	Improved motor activity	Ricceri et al. (2011)	–
Choline	ACh	Mecp2*tm1.1Bird*	NT	Improved motor coordination, anxiety and social behavior	Chin et al. (2018)	–
D-NAC	Dendrimer-conjugated N-acetyl cysteine	Mecp2*tm1.1Bird*	NT	Improved overall health, limb clasping	Nance et al. (2017)	–
Desipramine	Norepinephrine reuptake inhibitor	Mecp2*tm1.1Bird*	Y	Improved breathing irregularities	Roux et al. (2007)	Y
Desipramine	Norepinephrine reuptake inhibitor	Mecp2*tm1.1Bird*	Y	Reduced apneas	Zanella et al. (2008)	Y
Clenbuterol	B2-adrenergic receptor agonist	Mecp2*tm1.1Bird*	Y	Improved motor coordination and breathing irregularities	Mellios et al. (2014)	–
D-cycloserine	D-alanine analog	Mecp2*tm1.1Jae*	NT	No improvement	Na et al. (2017)	–
Growth factor signaling						
CX546	Ampakine (BDNF)	Mecp2*tm1.1Jae*	NT	Improved breathing irregularity	Ogier et al. (2007)	–
Fingolimod	Sphingosine-1 phosphate receptor (BDNF)	Mecp2*tm1.1Bird*	NT	Improved motor activity	Deogracias et al. (2012)	Y

CPT157633	PTP1B inhibitor (BDNF)	Mecp2*tm1.1Bird*	Y	Reduced limb clasping, partially improved motor activity	Krishnan et al. (2015)	–
LM22A-4	TrkB agonist (BDNF)	Mecp2*tm1.1Jae*	NT	Improved breathing irregularity	Schmid et al. (2012)	–
7,8-DHF	TrkB agonist (BDNF)	Mecp2*tm1.1Jae*	Y	Improved motor activity and breathing irregularities	Johnson et al. (2012)	–
LM22A-4	TrkB agonist (BDNF)	Mecp2*tm1.1Jae*	NT	Reduced apneas	Kron et al. (2014)	–
LM22A-4	TrkB agonist (BDNF)	Mecp2*tm1.1Jae*	NT	Improved memory	Li et al. (2017)	–
IGF-1	IGF-1	Mecp2*tm1.1Jae*	Y	Improves motor activity, breathing irregularities, increased brain size	Tropea et al. (2009)	Y
PEG-IGF-1	Slow release IGF-1	Mecp2*tm1.1Bird*	Y	No improvement	Pitcher et al. (2013)	–
RhIGF01	Recombinant human IGF1-1	Mecp2*tm1.1Bird*	Y	Improves motor activity, breathing irregularities, social behavior and anxiety	Castro et al. (2014)	Y
Metabolism						
Diet restriction	Caloric deficit	Mecp2*tm1Hzo*	NT	Improved motor activity and anxiety	Mantis et al. (2009)	–
Statins	Cholesterol-lowering medication	Mecp2*tm1.1Bird* Mecp2*tm1.1Jae*	Y	Improved overall health, motor coordination, motor activity, serum lipids and liver lipids	Buchovecky et al. (2013)	Y
Dietary triheptanoin	Energy use (mitochondria)	Mecp2*tm1.1Jae*	Y	Improved motor coordination, social behavior, insulin sensitivity, metabolic homeostasis	Park et al. (2014)	Y
Trolox	Vitamin E derivative	Mecp2*tm1.1Bird*	NT	Blood glucose levels normalized, improved response to hypoxia	Janc et al. (2016)	
Corticosterone	Glucocorticoid activation	Mecp2*tm1.1Bird*	Decreased lifespan	Worsened motor activity	Braun et al. (2012)	–
Corticosterone	Glucocorticoid activation	Mecp2*tm1Hzo*	NT	Improved motor activity	De Filippis et al. (2013)	–
RU486	Glucocorticoid activation	Mecp2*tm1.1Bird*	No	Delayed progression of symptoms, improved motor activity	Braun et al. (2012)	–

Curcumin	Anti-oxidant, antiinflammatory	Mecp2*tm1.1Jae*	NT	NT	Panighini et al. (2013)	–
Insulin	Glucose signaling	Mecp2*tm1.1Bird*	Decreased lifespan	No improvement	Pitcher et al. (2013)	–
Other						
Zoledronic acid	Anti-osteoclastic	Mecp2*tm1.1Bird*	NT	Increased bone volume and connectivity	Shapiro et al. (2010)	–
Cannabidivarin P	Phytocannabinoid	Mecp2*tm1Hzo*	NT	Improved overall health, motor activity and social behavior	Vigli et al. (2018)	–
CNF1	RhoGTPase	Mecp2*tm1Hzo*	NT	Improved motor activity	De Filippis et al. (2012)	–
CNF1	RhoGTPase	Mecp2*tm1Hzo*	NT	Improved mitochondrial dysfunction and memory	De Filippis et al. (2015)	–
Non-pharmacological						
Enriched environment	Environmental modulation	Mecp2*tm1Pplt*	NT	Improved motor coordination	Kondo et al. (2008)	–
Enriched environment	Environmental modulation	Mecp2*tm1.1Jae*	NT	Improved motor activity	Nag et al. (2009)	–
Enriched environment	Environmental modulation	Mecp2*tm1.1Jae*	N	Improved motor coordination and activity	Lonetti et al. (2010)	–
Enriched environment	Environmental modulation	Mecp2*tm1Pplt*	NT	Reduced anxiety	Kondo et al. (2016)	–
Forniceal deep brain stimulation	Neural circuit stimulation	Mecp2*tm1.1Bird*	NT	Improved memory	Hao et al. (2015)	–
Bone marrow transplantation	Brain microglia repopulation	Mecp2*tm1.1Bird*	Y	Reduced apneas, improved breathing irregularities, improved locomotor activity	Derecki et al. (2012)	–

Otras publicaciones

Xiol C, **Vidal S,** Pascual-Alonso A, Blasco L, Brandi N, Pacheco P, Gerotina E, O'Callaghan M, Pineda M, Armstrong J; Rett Working Group. X chromosome inactivation does not necessarily determine the severity of the phenotype in Rett syndrome patients. *Scientific Reports*. 2019 Aug 19;9(1):11983

Pascual-Alonso A, Blasco L, **Vidal S**, Gean E, Rubio P, O'Callaghan M, Martínez-Monseny AF, Castells AA, Xiol C, Català V, Brandi N, Pacheco P, Ros C, Del Campo M, Guillén E, Ibañez S, Sánchez MJ, Lapunzina P, Nevado J, Santos F, Lloveras E, Ortigoza-Escobar JD, Tejada MI, Maortua H, Martínez F, Orellana C, Roselló M, Mesas MA, Obón M, Plaja A, Fernández-Ramos JA, Tizzano E, Marín R, Peña-Segura JL, Alcántara S, Armstrong J. Molecular characterisation of Spanish patients with MECP2 duplication syndrome. *Clin Genet*. 2020 Feb 11. doi: 10.1111/cge.13718. [Epub ahead of print] PubMed PMID: 32043567.