The Tousled-like kinases and their implications in cancer and neurodevelopmental disorders

Marina Villamor Payà

UNIVERSITAT DE BARCELONA

Universitat de Barcelona
Facultat de Farmàcia i Ciències de l'Alimentació
Programa de Doctorat en Biomedicina

The Tousled-like kinases and their implications in cancer and neurodevelopmental disorders

Marina Villamor Payà

2021

Universitat de Barcelona
Facultat de Farmàcia i Ciències de l'Alimentació
Programa de Doctorat en Biomedicina

The Tousled-like kinases and their implications in cancer and neurodevelopmental disorders

Memòria presentada per **Marina Villamor Payà**
per optar al títol de Doctor per la Universitat de Barcelona



Director	Doctoranda	Tutor
Dr. Travis H. Stracker	*Marina Villamor Payà*	*Dr. Albert Tauler Girona*

Marina Villamor Payà
2021

The research described in this thesis was performed from October 2016 until December 2020 at the Institute for Research in Biomedicine (IRB Barcelona), and supported by a grant of the Ministry of Science, Innovation and Universities (MCIU), reference BFU2015-68354-P. Marina Villamor Payà was supported by a predoctoral FPI fellowship from the Ministry of Science, Innovation and Universities (MCIU), reference BES-2016-076423.

Acknowledgements

It takes a village to finish a PhD, and words cannot express how thankful I am for all the people who have helped me get here and have been by my side during the past four and a half years.

First and foremost, I'd like to thank my thesis director **Travis** for the opportunity to join your lab. Thanks for all the scientific discussion, and for your open-door policy (now zoom-me-anytime policy?). Thank you for always being open to start new projects and new collaborations and for listening and letting us follow our ideas. Thank you for encouraging me to leave my comfy "IRB bubble" and ship me to the USA alongside our TLK2 mice. It was a great experience, and I know how fortunate I am to have had the opportunity to go. Thank you also for being a true mentor and allowing me to participate in activities not so directly related to the science, and for always being available to discuss career options. Congrats again for your well-deserved position at the NCI, I am excited to see you succeed first-hand from there. And lastly, thank you for surrounding yourself and the lab with good people, it made for a very nice work environment I would not have changed for anything.

Thank you to the **lab(less) members**, Lluís, Ilaria, María, Isabel, Chithran, Michael (and Josep). I feel so lucky to have worked with you, you've been an immeasurable support through the whole PhD but specially during this last super weird year.

María, mi minion. Dicen que supervisar a un estudiante es una lotería y contigo gané el gordo. Somos un gran equipo tanto dentro como fuera del lab (cierrapisos) y echo de menos pasar todo el día contigo, aunque te tenga al teléfono cada dos por tres :P. **Josep**, the adopted child of the lab. No sé què hauria fet sense el teu bullying constant aquests anys. Gràcies per fer-me riure tant, pels cotilleos, pels teus consells de vida i per sempre estar allà. No tant gràcies per la misteriosa desaparició constant de xocolata i cafès... ;) **Ilaria**, thank you for your positive energy and big smile every morning, for always being ready to help others, for your constant scientific input and for your hard work. **Lluíiiiiiiiiiis** gràcies per la teva ajuda infinita amb els ratolins i els seus kikis, i per sempre estar on top de tot. Sense tu el lab no hauria funcionat! **Michael**, I am so glad to have gotten to know that under those huge muscles there's also a huge heart (and cilia) ;) **Isabel**, gracias por compartir tu conocimiento con nosotros,

and **Chithran**, I can't imagine how calm you must be without Josep & Mikel annoying all the time ;) I wish you all the best in your PhD. Thanks also to old members of the lab, **Andrea**, for the ice-cream and the Taylor Swift dance parties, i **Sandra**, per ser el millor exemple de feina ben feta.

Thank you **Anupriya** and **Seb** for taking me into your labs for my research stays abroad. My months in Portland and Munich have been very enriching at all the levels.

Als members del meu **thesis advisory committee**, Ángel Nebreda, Manuel Serrano i Joan Roig, gràcies per seguir la meva feina I donar-me input en cada reunió. Y Ángel, muchas gracias por adoptarme en tu lab aunque solo sea en papel estos últimos meses.

Maria José, gracias por toda tu ayuda con "nuestros niños" y por enseñarme todo lo que sé de neurodesarollo. Sin ti el proyecto de cerebritos no habría sido posible.

Gràcies a les IRB, PCB i UB facilities, per tota l'ajuda i per la vostra contribució en els nostres projectes al llarg d'aquests anys: **MS/Proteomics** (Marta, Marina, Gianluca) per l'ajuda en els mil i un BioIDs i per participar tan activament als proteomics data club, **Functional Genomics** (Nacho i David) per preparar les llibreries i arrays, **Histopatologia** (Neus, Mònica, Anaïs), **Citometria** (Jaume i Sonia) i a la Asun (Estabulari) per sempre ajudar-me amb el papeleo infinit. Gràcies a la **Biostatistics / Bioinformatics unit** (Camille, Toni) però sobretot a l'**Adrià**, per la teva paciència i per sempre trobar la millor manera d'analitzar els experiments i donar resposta a les meves preguntes.

Gràcies **Clara**, per deixar-me formar part de l'ENABLE 2018. Va ser una experiència molt bonica i vam aprendre molt: a treballar en equip, a delegar, a arribar a acords, a fer front a imprevistos d'últim minut. **Lorena** i **Enric**, no hauria volgut compartir aquesta aventura amb ningú altre. And the rest of **SOC members**, it was so much fun working with and learning from all of you. The 3 days of the symposium flew by and it was so rewarding to see the fruit of our hard work.

Leyre, gracias por estar siempre allí. Gracias por tu ayuda en el IRB, ayudándonos con cualquier papeleo y burocracias diversas des de las entrevistas hasta el final. Pero sobretodo gracias por tu amistad fuera del IRB, por los retreats a Holanda, los viajes improvisados, los planes de "tranquis" y los conciertos de Izal.

Cadiz-Pirineos team, **Craig**, **Leyre**, **Cris**, **Enric**, **Helena**, **Ricardo**, **Jürgen**, gracias chicos por las risas, los findes en Sallent "comiendo como gordis, con episodios conejo incluidos. La semana en Cadiz-Tarifa. Gracias Riki por tus manjares y por descubrir-nos la costa portuguesa, los juegos de palabras de Craig y Kekah, els smoothies healthies de l'Helena, la organización impecable de Leyre. Espero que podamos vernos pronto (y no por zoom, por favor).

Edu, **Sara**, **Joel**, **Isita**, **Lorena** i **Enric**, vam començar aquesta aventura junts i ara un parell ja són doctors i la resta estem a punt de ser-ho. Gràcies per compartir les alegries i les penes aquests anys. E Isa, gracias por enseñarnos que Valladolid en invierno mola ;)

Eugenia, **Clara**, **Iván**, **Leire**, **Amanda**. Thanks for making me feel like home, 10.000 km away from home. I can't imagine what my months in Portland would have been without you, our hikes to chase waterfalls and our Sunday night Spanish tortillas. Gracias, **Eugenia**, por adoptarme incluso antes de conocerme, i **Clarita** (o millor dit perleta del mediterrani) per les trucades a qualsevol hora i les nits veient Merlí. **Amanda**, mon pâtissier de macarons préférée, I miss our cooking sessions. Thank you, **Derek** and **Sophie**, for taking me to do all the American things and for trusting me with your car ;)

Raquel, **Montero**, **Marina**, dankeschön por hacer mis corona-meses en München mucho más llevaderos :) and **Bruno**, thank you for letting me be Leo's auntie and for embarking on this Graphit adventure with Reda and me. I am so excited to see how far we make it!

Nayeli, **Pieter**, **Charlotte**, **Christina**, **Virginia**, I love remembering our careless and naïve Erasmus days and the fun we had. But I specially love meeting you every now and then somewhere in the world and see how far we've come.

Sylvia, it has been so fun to see our parallel lives unfold over the years. Thank you for being my person and for understanding me. I am beyond excited to be sharing, again, our American adventures ;)

Adri, sempre seràs el petit de la casa, i m'encanta saber que sempre tindré assessorament legal gratis si la lio parda ;) **Rosa**, **Albert**, **Sergi**, **Dami**, **Marc**, **Eloy**, **Maria**, gràcies per fer-nos grans junts i per totes les festes, escapades i viatges al llarg dels anys. Espero que seguim celebrant dècades d'amistat.

Núria, gràcies per ser-hi sempre i per escoltar-me i entendre'm com ningú. Gràcies per les nostres aventures adolescents (i no tant adolescents), pels caps de setmana a Llafranc i els concerts de Joan Dausà.

Anna i **Diana**, gràcies per compartir amb mi aquells anys tan bonics que van ser la carrera, totes les teleco-bio-festes, i cases rurals, i viatges d'estiu, i les nits estudiant fins tard i els drames diversos, i el anar-nos a veure quan estàvem d'Erasmus. Gràcies també per tot el que ha vingut després: Londres, el PhD, el casament més bonic. No importa quan lluny vivim, espero que seguim trobant sempre la manera de continuar amb les nostres reunions improvisades i els Skypes a 3 bandes. Estic molt orgullosa de vosaltres.

Reda, thank you for your patience and love, and for feeling so close despite the distance. It was such a challenge at times, but we made it! This journey would have not been the same without you. Thank you for giving me mimos even when I give you grief, and for always staying calm when I freak out. Thank you also for always making me laugh, my hinte heltig cub. I can't wait to keep making memories together.

I per acabar, gràcies **mama**, gràcies **papa**, per l'amor incondicional i per donar-me les eines i la llibertat per escollir a cada pas el que volia fer. Sé que és tot un privilegi. Gràcies per recolzar sempre les meves decisions, encara que signifiquin viure molt lluny. I gràcies també pel privilegi de tenir tuppers de llenties de la iaia esperant-me cada cop que vinc a casa.

Table of contents

List of Figures ..... - 12 -
List of Tables ..... - 14 -
Summary ..... - 15 -
Resum ..... - 17 -
List of abbreviations ..... - 19 -
Introduction ..... - 23 -
Genomic Instability and the DNA damage response ..... - 24 -
Sensors of replication stress and double-strand breaks ..... - 24 -
The cell cycle and the DNA damage-induced checkpoints ..... - 26 -
Genome instability syndromes ..... - 28 -
Diseases of the central nervous system ..... - 28 -
Progeroid syndromes ..... - 29 -
Cancer ..... - 29 -
Replication stress in cancer ..... - 30 -
Neurodevelopment and neurodevelopmental disorders ..... - 31 -
Cortical development and different cell types in the brain ..... - 31 -
Neurogenesis ..... - 31 -
Gliogenesis ..... - 33 -
Microglia ..... - 34 -
Neurodevelopmental disorders ..... - 35 -
Genetic components and interactome of NDDs ..... - 35 -
Inflammation in NDDs ..... - 36 -
Microexon splicing defects in NDDs ..... - 37 -
The Tousled-like kinases (TLKs) ..... - 38 -
Discovery of the Tousled kinase and the TLKs ..... - 38 -
TLKs domains, structure and activation mechanism ..... - 39 -
Known substrates, interactors, regulators and functions of human TLKs ..... - 41 -
TLKs role in DNA replication, histone deposition, and chromatin maintenance ..... - 41 -
The role of TLKs in DNA repair ..... - 44 -
TLKs in mammalian development and disease ..... - 45 -
TLKs during murine development ..... - 45 -
TLKs in cancer ..... - 46 -
TLK inhibitors for cancer treatment ..... - 47 -
TLKs in neurodevelopmental disorders ..... - 49 -
Aims ..... - 51 -

Materials and methods ..... - 52 -
Animal studies ..... - 53 -
Genomic DNA isolation ..... - 53 -
Genotyping PCR ..... - 54 -
Tumor digestion ..... - 54 -
Flow cytometry ..... - 55 -
Sorting and Picoprofiling ..... - 55 -
Microarray ..... - 56 -
Tumor growth statistical analysis ..... - 58 -
Analysis of TCGA signatures ..... - 58 -
Tissue dissection for RNA-Seq ..... - 59 -
RNA extraction from mouse tissues for RNA-Seq ..... - 59 -
RNA-Seq ..... - 60 -
Differential analysis and biological enrichment ..... - 61 -
Alternative splicing analysis ..... - 61 -
Repetitive element analysis ..... - 62 -
Quantitative real-time PCR (qRT-PCR) ..... - 62 -
Protein extraction from mouse tissues for western blotting ..... - 63 -
Histopathology ..... - 63 -
Immunohistochemistry (IHC) ..... - 64 -
Cell lines and cell culture ..... - 65 -
Small interfering RNA (siRNA)-mediated knockdown ..... - 65 -
Genomic DNA extraction from human blood ..... - 66 -
Whole exome sequencing, array-based comparative genomic hybridization, prioritization, and variant calling ..... - 66 -
In silico prediction of missense variants impact ..... - 67 -
In silico modeling of TLK2 mutations ..... - 67 -
Single-cell gel electrophoresis (SCGE) ..... - 67 -
Site-directed mutagenesis ..... - 68 -
Transfection and Immunoprecipitation (IP) in mammalian cells ..... - 69 -
Western blotting analysis ..... - 69 -
In vitro kinase assays from cell lysates pull-downs ..... - 71 -
Alamar Blue viability assay ..... - 72 -
Immunofluorescence (IF) ..... - 73 -
Proximity-dependent biotin identification mass spectrometry (BioID-MS) ..... - 73 -
Small-scale BioID validation by western blot (BioID-WB) ..... - 75 -
Phospho-proteomics ..... - 75 -
Web Resources ..... - 77 -

Results ..... - 79 -
Chapter 1 – Tousled like kinases as a target in cancer therapy ..... - 80 -
TLK expression anti-correlates with immune signaling in many human cancers ..... - 82 -
Establishing a MMTV-PyMT mouse model to assess the role of TLK2 in breast cancer progression ..... - 84 -
TLK2 loss stalls primary tumor growth ..... - 87 -
TLK2 loss reduces lung metastasis incidence ..... - 89 -
Increased immune infiltration in TLK2 knockout tumors at early time points ..... - 92 -
Testing of small molecule inhibitors of TLK activity ..... - 98 -
Chapter 2 – TLK2 loss leads to inflammation in the developing brain ..... - 104 -
TLK2 knockout placentas do not exhibit increased inflammation ..... - 106 -
TLK2 knockout telencephalons have normal alternative splicing ..... - 109 -
TLK2 knockout embryos exhibit elevated transcription of inflammatory genes ..... - 113 -
Stromal cell components are upregulated in the TLK2 knockout telencephalons ..... - 116 -
Chapter 3 – Functional analysis and interactome characterization of TLK2 variants identified in MRD57 patients ..... - 120 -
Identification of 6 individuals with novel TLK2 variants ..... - 122 -
TLK2 mutations disrupt proper chromatin compaction ..... - 125 -
TLK2 missense mutations alter the activity and subcellular localization of the protein ..... - 126 -
The TLK2 interactome contains proteins involved in neurodevelopmental disorders ..... - 131 -
Missense mutations alter the proximal interactome of TLK2 ..... - 134 -
Discussion ..... - 143 -
TLKs as a potential cancer target to augment immune checkpoint blockade efficacy ..... - 144 -
Clinical implications of TLK2 mutations in neurodevelopment ..... - 148 -
TLK2 is part of a network of ASD susceptibility proteins ..... - 152 -
TLK2 in histone regulation and inflammation in the mouse embryonic brain ..... - 154 -
TLK2's proximal interactome and potential roles of TLKs in DNA damage repair and replication timing ..... - 158 -
Conclusions ..... - 163 -
Supplementary information ..... - 165 -
References 177

List of Figures

Figure I1. Scheme of the sensors of double-strand breaks and replication fork stalling and the DNA damage-induced checkpoints. ..... - 27 -
Figure I2. Development of the murine neocortex. ..... - 33 -
Figure I3. Domain architecture of human TLK1 and TLK2. ..... - 40 -
Figure I4. Model of human TLK activation. ..... - 41 -
Figure I5. TLKs have a role in histone deposition and genome and epigenome stability. ..... - 43 -
Figure I6. TLKs are amplified in cancer and in some cases high TLK2 expression correlate with reduced disease-free survival ..... - 47 -
Figure I7. Testing of TLK2 inhibitors. ..... - 49 -
Figure I8. TLK2 mutations in patients with NDDs ..... - 50 -
Figure R1. TLK expression correlates with suppressed immune signaling in human cancers. ..... - 84 -
Figure R2. Establishment of a MMTV-PyMT mouse model to assess the role of TLK2 in breast cancer progression. ..... - 86 -
Figure R3. TLK2 loss stalls primary tumor growth. ..... - 88 -
Figure R4. TLK2 loss reduces lung metastasis incidence. ..... - 91 -
Figure R5. TLK2 wild-type and knockout tumors have a similar TME composition. ..... - 93 -
Figure R6. TLK2 knockout tumors present with an early increased myeloid infiltration. ..... - 95 -
Figure R7. Early increased lymphoid infiltration in TLK2 knockout tumors. - 98 -
Figure R8. Testing of small molecule inhibitors of TLK activity. ..... - 100 -
Figure R10. TLK2 knockout placentas do not exhibit increased inflammation... - 108 -
Figure R11. Experimental workflow for telencephalon RNA-seq ..... - 111 -
Figure R12. TLK2 knockout telencephalons have normal alternative splicing. ... - 113 -
Figure R13. Innate immune response genes and histone variants are differentially regulated in TLK2 knockout telencephalons. ..... - 114 -
Figure R14. Increased expression of neuroinflammatory genes in TLK2 knockout embryos. ..... - 116 -
Figure R15. Stromal cell components are upregulated in the TLK2 knockout telencephalons. ..... - 118 -

Figure R16. Increased number of microglial cells in the cortex of *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice. ..... - 119 -
Figure R17. Identification of 6 individuals with novel TLK2 variants. ..... - 124 -
Figure R18. TLK2 mutations disrupt proper chromatin compaction. ..... - 125 -
Figure R19. Modelling TLK2 missense mutations. ..... - 128 -
Figure R20. TLK2 missense mutations alter the activity of the protein. ..... - 129 -
Figure R21. Missense mutations alter the subcellular localization of TLK2. - 131 -
Figure R22. The TLK2 interactome contains proteins involved in neurodevelopmental disorders. ..... - 133 -
Figure R23. Missense mutations alter the proximal interactome of TLK2 #1. ..... - 136 -
Figure R24. Missense mutations alter the proximal interactome of TLK2 #2. ..... - 137 -
Figure R25. Missense mutations alter the proximal interactome of TLK2 #3. ..... - 139 -
Figure R26. Validation of proximal interactors by BioID-WB. ..... - 140 -
Figure R27. Validation of proximal interactors by IP-WB. ..... - 141 -
Figure D1. Network graph representing a simplified scheme of the TLK2 interactome. ..... - 154 -
Supplementary figure 1. Phospho-proteomics analysis of hTERT-RPE-1 cells. . - 166 -
Supplementary figure 2. Increased expression of neuroinflammatory genes in TLK2 heterozygous embryos. ..... - 167 -
Supplementary figure 3. Repetitive element analysis of E13.5 placentas and E18.5 telencephalons. ..... - 168 -
Supplementary figure 4. Comparison of the proximal interactome of TLK1 and TLK2. ..... - 169 -

List of Tables

Table 1. Primers used for mouse genotyping. ..... - 54 -
Table 2. PCR program for *Tlk2*, *Cre*, *MMTV-PyMT*, *Sox2-Cre* alleles and Y chromosome..... - 54 -
Table 3. Antibodies used for flow cytometry and sorting. ..... - 56 -
Table 4. RNA-Seq sample information. ..... - 60 -
Table 5. qRT-PCR primers used in this study. ..... - 63 -
Table 6. siRNA sequences. ..... - 66 -
Table 7. Primers used for Sanger sequencing of patient variants. ..... - 67 -
Table 8. Site-directed mutagenesis primers. ..... - 68 -
Table 9. Primers used for Sanger sequencing. ..... - 69 -
Table 10. Primary antibodies used in this study. ..... - 70 -
Table 11. Secondary antibodies used in this study. ..... - 71 -
Table 12. Inhibitors tested in *in vitro* kinase assays ..... - 72 -
Table 13. Inhibitors used in viability assays. ..... - 72 -
Table 14. Web resources used in this thesis. ..... - 77 -
Supplementary table 1. TCGA cohort designation and number of patients per cohort (n). ..... - 170 -
Supplementary table 2. Comparison between TLK2 transcripts NM_001284333.2 and NM_006852.6. ..... - 172 -
Supplementary table 3. List of SFARI /DECIPHER /OMIM genes in the proximal interactome of wild-type and mutated TLK2. ..... - 173 -
Supplementary table 4. Number of overlapping genes between the different Mouse Cell Atlas cell populations 175

Summary

The Tousled like kinases 1 and 2 (TLK1 and TLK2) are evolutionarily conserved Ser/Thr kinases that are required for genome stability and normal development in numerous organisms. Both kinases contribute to histone deposition by targeting the histone chaperone ASF1 and are regulated by the DNA damage response. Their expression is ubiquitous and often elevated in cancer cells with suppressed immune gene expression. In cultured cells, TLK depletion causes replication stress, DNA damage and altered chromatin maintenance, eliciting an innate immune response dependent on the Alternative-lengthening of telomeres pathway. This suggested that TLKs would be a potential target in cancer therapy. In addition, TLKs have been linked to a distinct neurodevelopmental disorder (NDD) known as Mental Retardation Autosomal Dominant (MRD57). How TLKs contribute to these pathologies remains largely unknown. The aims of this thesis are to explore the therapeutic potential of targeting TLKs in cancer and understand how TLK2 deficiency contributes to NDDs. To this end, we have generated conditional mouse models where we can modulate TLK2 expression in both embryonic development and in developing breast cancer. We have also mimicked TLK2 mutations from MRD57 patients *in vitro* and addressed how they alter the normal kinase activity, localization and interactome of TLK2. We showed that depletion of TLK2 in already formed tumors stalls tumor growth and reduces lung metastases, and observed an early immune infiltration that could potentially sensitize these tumors to immune checkpoint blockade. Novel small molecule inhibitors towards TLKs were identified, indicating that they are inherently druggable targets. Consistent with what we observed in cancer, deletion of TLK2 in the telencephalon of mouse embryos caused elevated transcription of inflammatory genes and signatures of microglia and astrocytes. We also described 6 new cases of MRD57 and showed that patient-derived cell lines exhibited defects in chromatin compaction. Separately we identified two hyperactive MRD57 mutations, suggesting misregulation of TLK2 levels, rather than decreased activity, is likely the predisposing factor leading to NDD. Lastly, we showed in cells that TLK2 is in close proximity with a number of known autism-susceptibility proteins and proteins enriched in replication forks, and that these proximal interactions are altered by MRD57 mutations. Together, these results further suggest that TLKs may be a viable target for cancer therapy, as their loss leads to innate immune activation that may also influence neurodevelopment in MRD57 patients.

Resum

Les Tousled like kinases i and 2 (TLK1 i TLK2) són Ser-Thr quinases conservades evolutivament que són necessàries pel manteniment de l'estabilitat del genoma i pel desenvolupament normal de diversos organismes. Ambdues quinases contribueixen a la deposició d'histones mitjançant la regulació de la xaperona d'histones ASF1, i són regulades per les vies de resposta al dany de l'ADN. La seva expressió és ubiqua i sovint elevada en cèl·lules canceroses que presenten l'expressió de gens immunes suprimida. En cèl·lules cultivades, la depleció de les TLKs causa estrès replicatiu, dany en l'ADN i altera el manteniment de la cromatina, provocant una resposta immune innata depenent de la via d'allargament alternatiu dels telòmers. Això suggereix que les TLKs són una diana terapèutica potencial per al tractament del càncer. A més a més, les TLKs han estat vinculades a un trastorn del neurodesenvolupament (TND), anomenat Retard mental autosòmic dominant (MRD57). Com les TLKs contribueixen a aquestes patologies continua sent un misteri. Els objectius d'aquesta tesi són explorar el potencial terapèutic en càncer de les TLKs i entendre com el dèficit de TLK2 contribueix a TND. Per assolir aquests objectius, hem generat models condicionals de ratolí on podem modular l'expressió de TLK2 tant en desenvolupament embriònic com durant la progressió de tumors mamaris. També hem mimetitzat mutacions de pacients amb MRD57 *in vitro* i estudiat com alteren l'activitat quinasa, la localització i l'interactoma normal de TLK2. Hem demostrat que la depleció de TLK2 en tumors mamaris n'atura el creixement i redueix significativament les metàstasis als pulmons. També hem observat un augment en la infiltració de cèl·lules immunes als tumors que poden sensibilitzar-los a inhibidors dels punts de control immunitaris. Hem identificat nous inhibidors de molècules petites de les TLKs, reafirmant que són dianes drogables. Consistent amb el que hem vist en càncer, la supressió de TLK2 en el telencèfal d'embrions de ratolí causa un augment en la transcripció de gens inflamatoris i en signatures de micròglia i astròcits. Hem descrit 6 nous casos de MRD57 i demostrat que línies cel·lulars derivades d'aquests pacients presenten defectes en la compactació de la cromatina. Separadament hem identificat dues mutacions de MRD57 que són hiperactives, suggerint que la desregulació dels nivells de TLK2 predisposen al desenvolupament de TND. Per últim, hem demostrat en cèl·lules que TLK2 està a prop de diverses proteïnes que confereixen susceptibilitat a autisme, així com de proteïnes que localitzen a les forquilles de replicació. Aquestes interaccions es veuen alterades per les mutacions de MRD57. Conjuntament, aquests resultats

ens indiquen que les TLK2 són una diana viable pel tractament del càncer, ja que la seva pèrdua activa una resposta innata immunitària, i aquesta mateixa resposta sembla que també influeix en el neurodesenvolupament de pacients amb MRD57.

List of abbreviations

9-1-1	RAD9-RAD1-HUS1
^{32}P	Phosphorus-32
4-OHT	4-hydroxytamoxifen
A-T	Ataxia Telangectasia
A-T-LD	Ataxia Telenectasia-like diseases
aa	Aminoacid
ACMG	American College of Medical Genetics
ADHD	Attention deficit hyperactivity disorder
ALT	Alternative-lengthening of telomeres
APB	ALT-associated PML body
array-CGH	Array-based comparative genomic hybridization
AS	Alternative splicing
AS	Analog-sensitive
ASC	Autism Sequencing Consortium
ASD	Autism spectrum disorder
ASF1	Anti-silencing factor 1
ATAC	Assay for Transposase-Accessible Chromatin
ATM	Ataxia-telangiectasia mutated
ATR	Ataxia telangiectasia mutated and Rad3-related
ATRIP	ATR-interacting protein
ATRX	Alpha thalassemia / mental retardation syndrome X-linked
BBB	Blood-brain barrier
CBC	Complete blood count
CC	Coil-coiled
CDC25	Cell division cycle 25
CDK	Cyclin dependent kinase
cDNA	Complementary DNA
cGAS	Cyclic GMP–AMP synthase
CHD	Chromodomain Helicase DNA Binding Protein
CHK1	Checkpoint kinase 1
CHK2	Checkpoint kinase 2
CNS	Central nervous system
CP	Cortical plate
CR	Cajal-Retzius
CSF	Cerebrospinal fluid
CSR	Class switch recombination
CyTOF	Cytometry by time of flight
DAXX	Death domain-associated protein

DDR	DNA damage response
DECIPHER	*DatabasE* of genomiC varIation and Phenotype in Humans using Ensembl Resources
DNA	Deoxyribonucleic acid
DNM	*De novo* mutation
DNMT	DNA methyltransferase
DSB	Double-strand break
ECM	Extracellular matrix
ECTR	Extrachromosomal telomeric sequence
ER	Estrogen receptor
ER+	Estrogen receptor positive
ERT2	Mutant Estrogen Receptor
ERVs	endogenous retroviruses
FDA	Food and Drug Administration
FDR	False discovery rate
gDNA	genomic DNA
gnomAD	The Genome Aggregation Database
GO	Gene Ontology
GSEA	Gene set enrichment analysis
Gx-phase	Gap x-phase
HET	Heterozygous
ICB	Immune checkpoint blockade
ID	Intellectual disability
IF	Immunofluorescence
IFN	Interferon
IL	Interleukin
INM	Interkinetic nuclear migration
IP	Intermediate precursor
IQ	Intelligence quotient
IR	Ionizing radiation
IZ	Intermediate zone
KO	Knock-out
LCLs	Lymphoblastoid cells
LOF	Loss-of-function
LPS	Lipopolysaccharide
MBP	Myelin basic protein
MEF	Mouse embryonic flibrblast
mESC	Mouse embryonic stem cells
MIA	M1aternal immune activation
MMTV	Mouse mammary tumor virus
MRD57	Mental Retardation Autosomal Dominant 57

MRE11	Meiotic recombination 11
MRN	MRE11-RAD50-NBS1
MS	Mass spectrometry
MZ	Marginal zone
NBS	Nijmegen breakage syndrome
NBSLD	Nijmegen breakage syndrome-like disorder
NDD	Neurodevelopment disorder
NEC	Neuroepithelial cell
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NF1A	Nuclear factor 1A
NLS	Nuclear localization signal
NMD	Nonsense mediated mRNA decay pathway
nt	Nucleotide
OCD	obsessive-compulsive behaviour
OL	Oligodendrocyte
OPC	Oligodendrocyte precursor
P	Phospho
PARP	Poly-(adenosine 5'-diphosphate-ribose) polymerase
PCA	Principal component analysis
PD-1	Programmed cell death protein 1
PD-L1	Programmed death-ligand 1
PE	Paired-end
Pen-Strep	Penicillin-Streptomycin
PKD	Protein Kinase Domain
PML	Promyelocytic leukemia protein
PMZ	Promazine hydrochloride
PP	Preplate
PP1	Protein phosphatase 1
PPH	Perphenazine
PPI	Protein-protein interaction
PR	Progesterone receptor
PSI	Percentage of spliced-in
PyMT	Polyoma middle T antigen
RBP	RNA-binding protein
RCAS	Replication-competent avian sarcoma-leukosis virus
RGC	Radial glial cell
RIF1	Replication timing regulatory factor 1
RIN	RNA integrity number
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species

RPA	Replication protein A
rRNA	ribosomal RNA
RS	Replication stress
RSR	Replication stress response
SE	Single-end
SFARI	Simon's Foundation Autism Research Initiative
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
SP	Subplate
SSB	Single strand break
ssDNA	Single-stranded DNA
STING	Stimulator of IFN gene
SVZ	Subventricular zone
TAM	Tumor-associated macrophage
TBK1	TANK-binding kinase 1
TCGA	The Cancer Genome Atlas
TCR	T-cell receptor
THD	Thioridazine hydrochloride
TLK	Tousled-like kinase
TLR	Toll-like receptor
TME	Tumor microenvironment
TNBC	Triple negative breast cancer
TNFα	tumor necrosis factor alpha
TOPBP1	Topoisomerase-binding protein 1
UbCre	UbiquitinC-Cre
UV	Ultraviolet
VPA	Valproic acid
VZ	Ventricular zone
WES	Whole exome sequencing
WT	Wild-type
ΔPSI	Differential PSI

Introduction

Genomic Instability and the DNA damage response

The integrity of the genome is essential for organismal viability and reproduction. Cells are under a constant threat of DNA damage caused by a wide range of environmental and endogenous mutagenic agents, such as ultraviolet (UV) light, ionizing radiation (IR), pollutants, pharmaceuticals, reactive oxygen species (ROS) and byproducts of normal DNA replication and cell metabolism. DNA breaks also occur in a programmed manner to generate genetic diversity, for example, in developing B- and T- lymphocytes to increase the immunoglobulin and T-cell receptor (TCR) repertoire, as well as during meiosis where breaks are generated to stimulate recombination (Jackson and Bartek, 2009).

If unrepaired, DNA lesions can be cytotoxic or give rise to mutations that can ultimately lead to disease. Through evolution, cells have developed a battery of molecular mechanisms to protect and repair the DNA to ensure genome integrity and its faithful transmission to the offspring (Felter BC, 2016). These include systems, collectively known as the DNA damage response (DDR), that quickly detect a damaged lesion, signal its presence and mediate DNA damage repair or other cell fate decisions (Ciccia and Elledge, 2010). The DDR orchestrates a complex signaling network that controls cell cycle checkpoints, coordinates DNA replication with DNA repair and ultimately controls cell fate, to prevent the accumulation of genomic instability.

Genomic instability refers to the appearance of a high frequency of mutations in a genome. These mutations can include changes in a few nucleotides (point mutations, insertions/deletions), chromosomal breaks or rearrangements, or gain or loss of whole chromosomes (aneuploidy) (Lengauer et al., 1998).

Sensors of replication stress and double-strand breaks

The DDR provides spatio-temporal support for DNA sensing and repair by controlling cell cycle progression (Blackford and Jackson, 2017). Each type of DNA lesion is recognized by a specific sensor protein, that recruits repair proteins to the site of damage that will later signal downstream.

One of the major sensors of DNA double-strand breaks (DSBs) in Gap (G1)-phase is the MRE11-RAD50-NBS1 (MRN) complex, composed of Meiotic recombination 11 (MRE11), RAD50, and Nijmegen breakage syndrome (NBS1) proteins. The MRN complex rapidly recognizes DSBs and recruits and activates the Ataxia-telangiectasia mutated (ATM) protein kinase at the site of damage. Once activated, ATM starts a cascade of phosphorylation events, further

phosphorylating the MRN complex and its downstream substrate Checkpoint kinase 2 (CHK2) (Figure I1), as well as potentially hundreds of other substrates (Stracker and Petrini, 2011, Matsuoka et al., 2007).

Replication stress (RS) is a state defined by the slowing or stalling of replication fork progression and/or DNA synthesis, increased single-stranded DNA (ssDNA) accumulation and activation of the DDR. RS arises when replication forks are challenged during S-phase by unrepaired DNA lesions, lack of deoxynucleotides, secondary DNA structures or replication-transcription conflicts. As a result, the polymerase stalls and the replicative helicase uncouples, continuing to unwind the parental DNA producing ssDNA (Pacek and Walter, 2004, Byun et al., 2005, Zeman and Cimprich, 2014). This ssDNA is then bound by the tripartite complex Replication protein A (RPA). The ssDNA-RPA complex, along with the 9-1-1 (RAD9-RAD1-HUS1) DNA clamp, serves as a platform to signal the accumulation of ssDNA and recruit and activate replication stress response (RSR) proteins, including the kinase Ataxia telangiectasia mutated and Rad3-related (ATR) and its interaction partners ATR-interacting protein (ATRIP) and topoisomerase-binding protein 1 (TOPBP1) (Figure I1) (Zou and Elledge, 2003, Nam and Cortez, 2011). ATR activation and subsequent phosphorylation (p-) of the Checkpoint kinase 1 (CHK1) serves to stabilize arrested forks, suppress late origin firing, and activate repair machinery to faithfully complete DNA replication under stress conditions (Cimprich and Cortez, 2008). Furthermore, poly-(adenosine 5'-diphosphate-ribose) polymerase (PARP) activity is necessary for efficient CHK1 retention at forks, S-phase checkpoint activation and restart of stalled forks, to prevent from a prolonged fork arrest, which poses a risk for fork collapse and the generation of DSBs (Bryant et al., 2009, Min et al., 2013).

The ATR-CHK1 pathway is also activated in response to DSBs by the process of break resection. DSB undergo 5'-3' nucleolytic degradation that is initiated by the MRN complex protein MRE11, and is carried out by the action of several nucleases including MRE11, EXO1 and DNA2. This results in 3' ssDNA overhangs that are bound by RPA and similarly to stalled replication forks trigger ATR-CHK1 signaling (Figure I1) (Jazayeri et al., 2006). These single-stranded overhangs also play an important role in the choice of DSB DNA repair pathway (Huertas, 2010, Stracker and Petrini, 2011).

Once activated, ATM and ATR amplify a signal transduction cascade critical to activate cell cycle checkpoints to prevent cells from progressing through the cell cycle with damaged DNA. This arrest allows for DNA repair to take place before

re-entering the cell cycle. However if the damage is unrepairable, cells can activate other cell fate programs, such as senescence or apoptosis, that prevent the propogation of damage.

The cell cycle and the DNA damage-induced checkpoints

The cell cycle is the processes through which a proliferating cell accumulates resources, replicates its DNA and divides into two identical daughter cells. The cell cycle is divided into four phases: Gap 1 (G1), DNA synthesis (S), Gap 2 (G2), and Mitosis (M).

The progression through the cell cycle is tightly regulated by different cyclin and cyclin dependent kinase (CDKs) complexes, that are expressed and activated in a stepwise manner at different stages of the cell cycle (Malumbres, 2014). Particularly, in early G1 phase expression and activation of the cyclin D/CDK4/6 complex initiates a new cell cycle (Bertoli et al., 2013). The G1/S transition requires the cyclin E/CDK2 complex; cyclin A/CDK2 allow DNA replication during S phase, while cyclin B/ CDK1 allow for the G2/M transition and the initiation of mitosis (Lim and Kaldis, 2013). Cyclin/CDK complexes are regulated at different levels, including by phosphorylation. Cell division cycle 25 (CDC25) phosphatases dephosphorylate and activate CDKs, while WEE1 kinase phosphorylates CDK1 and CDK2 and inactivates them (Figure I1) (Shaltiel et al., 2015).

In response to DNA damage, three cell cycle checkpoints can be activated, namely the G1/S, intra-S, and G2/M checkpoints. The signals that link DNA damage sensing to checkpoint activation rely on the ATM-CHK2 and ATR-CHK1 pathways depending on the stimuli (Jazayeri et al., 2006). CHK1 and CHK2 kinases can stop the cell cycle by phosphorylating and inhibiting the CDC25 family phosphatases to suppress CDK function. Inhibition of cyclin E/CDK2 activates the G1/S checkpoint, inhibition of cyclin A/CDK2 activates the intra-S phase chekcpoint and inhibition of cyclin B/CDK1 stalls G2/M progression (Yata and Esashi, 2009). ATM and CHK2 can also phosphorylate the tumor suppressor p53 and the p53-specific ubiquitin ligase MDM2. In unstressed conditions, MDM2 constantly monoubiquitinates p53 and mediates its degradation through the proteasome (Figure I1). Once p53 and MDM2 are phosphorylated, their interaction is disrupted and p53 stabilized (Moll and Petrenko, 2003). P53 can then act as a transcription factor and promote the expression of the CDK inhibitor p21, which can prevent the G1/S transition and

cause G1 arrest and cell cycle exit; as well as genes involved in other cell fate programs such as apoptosis and senescence (Senturk and Manfredi, 2013).

The G2/M checkpoint can be activated by WEE1 phosphorylation and inhibition of CDK1. Polo-like kinase 1 (PLK1), that plays a key role in mitotic progression, becomes indispensable for G2 checkpoint recovery, as it can phosphorylate and activate the CDC25 phosphatases, as well as phosphorylate WEE1 and induce its degradation, contributing to both the activation of cyclin B/CDK1 and re-entry to mitosis (Shaltiel et al., 2015).

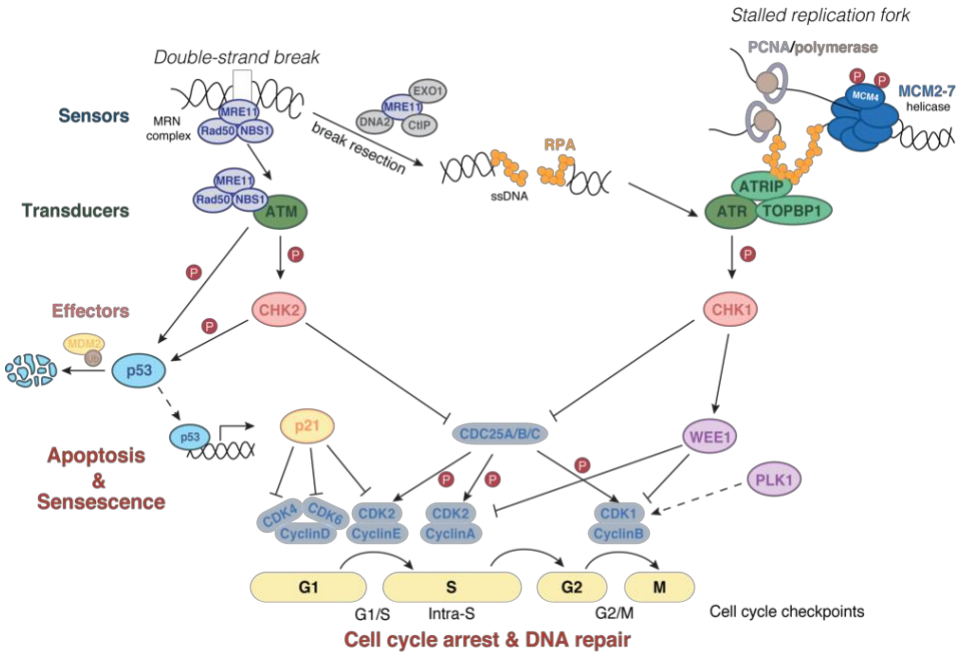


Figure I1. Scheme of the sensors of double-strand breaks and replication fork stalling and the DNA damage-induced checkpoints.

DSBs and stalled replication forks are sensed by different mechanisms that activate ATM and ATR, respectively. This triggers a phosphorylation cascade that culminates in the regulation of CDK-cyclins and the control of the cell cycle by activation of the checkpoints, that allow for the repair of the DNA before re-entering the cell cycle, or end up in apoptosis or senescence of the cell. Adapted from (Bouwman and Jonkers, 2012).

Genome instability syndromes

A defective DDR is associated with disease in the context of rare chromosome instability syndromes. Chromosome instability syndromes are caused by inactivating mutations in essential DNA repair/DDR-related genes and are characterized by an accumulation of DNA lesions that cause a wide spectrum of pathologies, including neurological abnormalities, including microcephaly and neurodegeneration, premature aging, infertility and predisposition to various types of cancer (Jackson and Bartek, 2009, Terabayashi and Hanada, 2018).

Diseases of the central nervous system

Patients with genome instability syndromes often present with immunodeficiency and infertility, that can be explained by the faulty repair of DSBs during V(D)J recombination, class switch recombination (CSR) and meiotic recombination. The accumulation of DNA lesions, as well as impaired signaling responses, due to defects in the DDR during brain development can result in neurodegeneration and microcephaly (Stracker et al., 2013). It however remains unclear precisely why these patients present with neurological pathologies, as there are not known physiological DSBs generated during neurogenesis. Some hypotheses include an accumulation of reactive oxygen species (ROS) due to the high oxygen consumption in neurons, resulting in oxidative stress and subsequent DNA damage (Jackson and Bartek, 2009), the fact that damaged neurons cannot be replaced during adulthood, or that neurons depend on the error-prone NHEJ pathway to repair DSBs because they are in G0 phase (Stracker et al., 2013, Terabayashi and Hanada, 2018). Recent work has shown that neuron specific transcription is a likely source of single strand breaks (SSBs) during development of the brain (Reid et al., 2021, Wu et al., 2021).

Central nervous system (CNS) pathologies are a common feature of diseases caused by mutations in members of the ATM and ATR-dependent DDR. These patients present with significant phenotypic heterogeneity. Depending on the particular gene, mutation, and patient, they can result in either microcephaly or neurodegeneration. Patients with mutations that disrupt the function of ATM, the main transducing kinase upon DSBs, suffer from Ataxia telangiectasia (A-T). A-T is characterized by ataxia and progressive neurodegeneration, caused by the atrophy of the cerebellum and the loss of Purkinje and granule cells. Other pathologies associated with A-T include severe immunodeficiency, sterility, increased radiosensitivity and proneness to cancer development, especially lymphomas and leukemias (Barlow et al., 1996, McKinnon, 2012,

Rothblum-Oviatt et al., 2016). Mutations in the components of the MRN complex, that senses DSBs upstream of ATM and plays enzymatic and signaling roles in the DDR, cause disorders with similar neuropathologies to A-T: Ataxia Telangiectasia -like diseases (A-T-LD) (*MRE11* mutation), Nijmegen breakage syndrome (NBS) (*NBS1* mutation), and NBS-like disorder (NBSLD) (*RAD50* mutation) (Stracker and Petrini, 2011). Cells derived from these patients present with increased levels of chromosomal instability, show high sensitivity to DSBs, and exhibit impaired checkpoint activation (McKinnon, 2012).

On the other hand, patients with hypomorphic mutations in ATR, the main transducer of the DDR upon replication stress, suffer from Seckel syndrome. Seckel syndrome is characterized by severe intrauterine growth delay, proportionate dwarfism, microcephaly, mental retardation and skeletal and brain abnormalities (Majewski and Goecke, 1982, Cherian, 2004, Murga et al., 2009). Several other genes implicated in DDR have been found in patients with Seckel-like syndrome, including the ATR partner *ATRIP*, and the resection enzymes *CTIP* and *DNA2*, as well as genes implicated in centrosome biogenesis (*CEP63*, *CEP152*, *CEP215*, or *PCNT*) (Kalay et al., 2011, Qvist et al., 2011, Sir et al., 2011, Yigit et al., 2015).

Progeroid syndromes

Mutations in DDR and DNA repair genes also cause hereditary progeroid syndromes, that are characterized by accelerated aging, growth retardation, short stature, and increased cancer predisposition. Among them, mutations in RecQ family helicases, that are important for the resolution of replication fork structures such as G-quadruplexes, cause Werner syndrome (*WRN* mutations) (Yu et al., 1996), Bloom syndrome (*BLM* mutations), and Rothmund-Thomson syndrome (*RECQL4* mutations) (Suhasini and Brosh, 2013). While poorly understood at the mechanistic level, stem cell aging has been implicated in the pathological outcomes (Cheung et al., 2015).

Cancer

Genomic instability is a inherent characteristic of most cancers and has been proposed as one of the earliest events in the evolution of a normal cell into a cancer cell (Hanahan and Weinberg, 2011). Hereditary cancers are characterized either by microsatellite instability or chromosomal instability (CIN) that are caused by germline mutations in DNA repair and DDR-related genes (Negrini et

al., 2010). The most well-known examples are mutations in *BRCA1* or *BRCA2* genes, that are involved in DSB repair via the homologous recombination (HR) pathway (Moynahan et al., 1999, Moynahan et al., 2001), and account for 25% of cases of hereditary breast and ovarian cancer (Miki et al., 1994, Wooster et al., 1995). Other frequent mutations are in *MSH2* and *MLH1* genes, that are involved in the mismatch repair (MMR) pathway, and cause Lynch syndrome, predispose to a wide range of tumors, including colon cancer (Fishel et al., 1993, Papadopoulos et al., 1994, Lynch et al., 2009).

Replication stress in cancer

Somatic mutations in DDR genes are also observed in several cancer types. It is now well established that the overexpression or activation of oncogenes involved in sustained cellular proliferation, such as *MYC* or *CCNE1,* during early tumorigenesis can drive unscheduled DNA replication (Gorgoulis et al., 2005, Bartkova et al., 2005). This unscheduled replication generates an environment where nucleotides, histones, and replication factors are limiting, ultimately leading to fork stalling and collapse, and ssDNA exposure. The ssDNA tracts are substrates for endonuclease cleavage, that generate DSBs, preferentially at fragile sites, that in turn activate the DDR (Macheret and Halazonetis, 2015). The first evidence for this oncogene-induced DNA replication stress model (Halazonetis et al., 2008) was the demonstration that early stage-tumors exhibit high levels of DNA damage signaling markers (such as phospho (P)-ATM and P-P53). This implies that the DDR acts as an inducible barrier to cancer progression by initiating programs that lead to cancer cell apoptosis or senescence (Gorgoulis et al., 2005, Bartkova et al., 2005). Excessive genome instability and subsequent DDR activation impair growth of the tumor, setting a selective pressure for tumors to bypass this barrier in order to continue proliferating in the presence of genomic instability. The most common way to do so is for tumors to inactivate specific components of the DDR, such as the tumor suppressor P53 or the ATM kinase. As a reference, *P53* is the most commonly mutated gene in cancer (more than 40% of all cancers), while *ATM* is at the top 20 pan-cancer mutated genes (Macheret and Halazonetis, 2015). Lack of P53 or ATM allow cancer cells to sustain proliferation in the presence of DNA damage accompanied by an accumulation of RS and further genomic instability, without proper activation of the apoptosis or senescence pathways. It also allows for the generation of additional mutations that favor cancer development, like those that further activate oncogenes or inactivate tumor suppressor genes, and for the generation

of chromosomal rearrangements driven by replication-transcription conflicts (Macheret and Halazonetis, 2015, Kotsantis et al., 2018).

Interestingly, deleterious mutations in RSR genes (such as ATR and CHK1) are rarely observed in cancer. This suggests that cancer cells heavily rely on the ATR/CHK1 axis to maintain chromosomal instability (CIN) and RS at subtoxic levels, representing a potential target for intervention (Forment and O'Connor, 2018). Inhibition of ATR or CHK1 in mice lacking ATM or p53, impairs tumor growth in MYC-overexpressing lymphomas, as well is in a number of other experimental models. In the context of MYC, the RS levels are so high when ATR-CHK1 is inhibited that it becomes toxic for the cancer cells (Murga et al., 2011). These results highlight the potential of inhibiting the ATR/CHK1 pathway to induce RS as a therapy, particularly in oncogene addicted tumors (Toledo et al., 2011). Currently, several ATR and CHK1 inhibitors are at different stages of development with some already in clinical trials (Lecona and Fernandez-Capetillo, 2018).

Neurodevelopment and neurodevelopmental disorders

Cortical development and different cell types in the brain

Neurogenesis

Mouse brain development starts around embryonic day (E) 7 (E7) with the formation of the neural plate, which consists in a thickened region of the embryonic ectoderm. Around E8.5, the neural plate folds to form the neural tube, that will become the future central nervous system (CNS). Around E9 the neural tube closes to form the vesicles that will later develop into the cortex. The neocortex emerges from the most rostral part of the neural tube. Neurogenesis in the developing neocortex extends from E11 to E18, followed by a period of gliogenesis that terminates after birth (Gotz and Huttner, 2005, Chen et al., 2017). Our description of embryonic neurogenesis focuses on the mouse dorsal neocortex as does the majority of the available literature.

All neocortical cells, except microglia, derive from the earliest progenitors of the cortex named neuroepithelial cells (NECs). NECs are organized in a

pseudostratified neuroepithelium in the neural tube (Gotz and Huttner, 2005). Initially, NECs expand their pool by undergoing successive symmetric divisions (McConnell, 1995) (Figure I2). Later, these cells divide asymmetrically giving rise to pluripotent radial glial cells (RGCs) (Gotz and Huttner, 2005). Both NECs and RGCs are located at the ventricular zone (VZ) and present with apical-basal polarity and interkinetic nuclear migration (INM), with mitosis taking place in the apical side (Miyata, 2008).

At the onset of neurogenesis, RGCs begin to divide asymmetrically producing a daughter RGC and either a differentiated neuron (neurogenic division) or a neurogenic-restricted progenitor named intermediate precursor (IP, proliferative division). IPs detach from the VZ losing the apico-basal polarity and migrate to a more basal region, the subventricular zone (SVZ). At the SVZ most IPs undergo symmetrical divisions and give rise to two identical excitatory glutamatergic neurons (Noctor et al., 2004, Sessa et al., 2008). Inhibitory GABAergic interneurons are exclusively derived from RGCs of the ventral telencephalon and migrate long distances until they reach their final position (Wonders and Anderson, 2006).

The first dorsal neurons generated are Cajal-Retzius (CR) and subplate neurons that form a transient layer called the preplate (PP). CR neurons migrate to the surface of the cortex forming the superficial marginal zone (MZ), while the rest of the PP evolves into the subplate (SP) (Angevine and Sidman, 1961, Luskin and Shatz, 1985).

The neurons produced in the VZ and SVZ migrate outwards and accumulate in the cortical plate (CP), which develops in between the MZ and the SP in an inside-out manner. The first neurons generated occupy the lower layers of the CP while the latest neurons migrate to form the most external layers (Malatesta et al., 2000, Noctor et al., 2001) (Figure I2). This results in 6 highly-organized cortical layers that contain both neuronal and glial types. The neurons located in each layer express different transcription factors, have specific morphology and send their projections to different parts of the brain, ultimately serving different functions (Reviewed in Molyneaux et al., 2007).

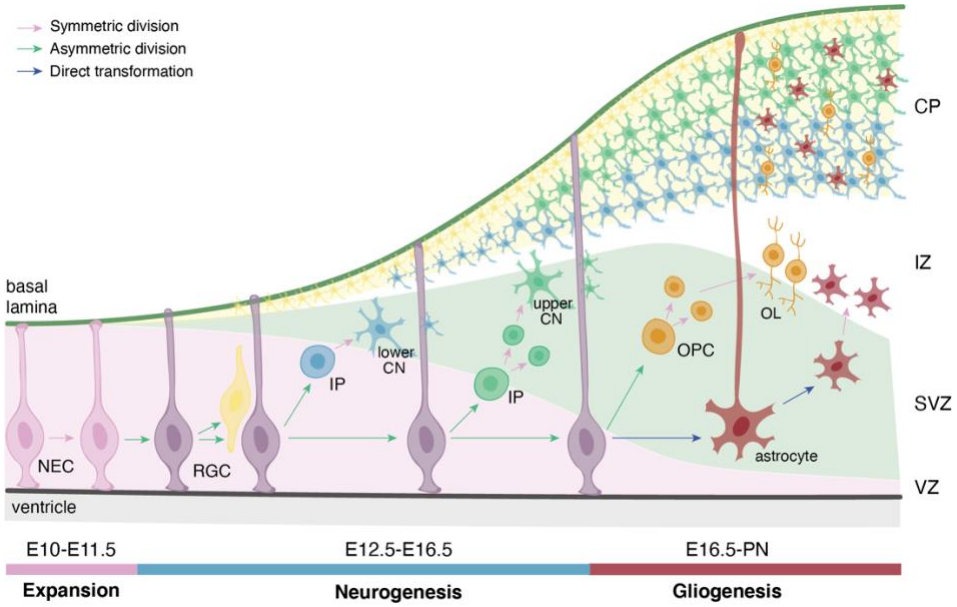


Figure I2. Development of the murine neocortex.

Glutamatergic neurons and cortical glial cells are generated in a sequential manner through symmetric and asymmetric divisions from common progenitor cells named neuroepithelial cells (NECs). During early stages of development, NECs expand their pool by symmetric divisions, and then start generating radial glial cells (RGCs). RGCs divide asymmetrically to give rise to neurons or intermediate progenitors (IP). The RGC and IP-derived cortical neurons (CN) are generated in an inside-out manner, where the lower neuronal layers are produced first. At around day E18, there is a neurogenesis to gliogenesis switch, where the remaining RGCs produce oligodendrocyte precursors (OPCs) through asymmetric division or directly transform into astrocytes. VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; SP, subplate; CP, cortical plate; MZ, marginal zone; OL, oligodendrocyte. Adapted from (Agirman et al., 2017).

Gliogenesis

Towards the end of embryogenesis (around E18), neurogenesis is suppressed and gliogenesis starts. The remaining RGs of the dorsal telencephalon stop producing neurons and start generating macroglial cells: first oligodendrocyte precursor cells (OPCs), that will later mature into olygodendrocytes (OL), and then astrocytes (Schmechel and Rakic, 1979, Voigt, 1989, Misson et al., 1991, Chanas-Sacre et al., 2000).

This neurogenic to gliogenic switch requires the activation of the Notch and Jak/Stat pathways, as well as the transcription factors SOX9 and Nuclear factor 1A (NF1A), which leads to the demethylation and expression of astrocyte specific genes (He et al., 2005, Kang et al., 2012, Namihira et al., 2009).

At birth, glial cells comprise only a 3% of the cells in the mouse cortex. Unlike the post-mitotic neurons that undergo terminal mitosis, astrocytes continue to divide locally after they have migrated (Ge et al., 2012), and glial cells (including microglia) end up representing 50% of the CNS volume through adulthood (Bandeira et al., 2009, and reviewed in Ohtsuka and Kageyama, 2019).

In the adult brain, astrocytes secrete substances that regulate neuronal differentiation, migration and survival, as well as the formation, plasticity and maintenance of synapses. Furthermore, astrocytes contribute to the maintenance of the blood-brain barrier (BBB) and participate in mounting an inflammatory response in the brain upon an insult (Reviewed in Linnerbauer et al., 2020). Oligodendrocytes on the other hand, form and maintain the myelin shield that enwraps the neuronal axons in the CNS, allowing for a quick transmission of the neural impulses (Reviewed in Kuhn et al., 2019).

Microglia

Apart from oligodendrocytes and astrocytes, there is another important glial component in the brain that does not arise from NECs: microglial cells. Microglia are myeloid cells that emerge early in development during the first wave of primitive hematopoiesis (Ginhoux et al., 2010). Microglia originate from hematopoietic progenitors within the extraembryonic yolk sac, and invade the developing brain at around E8, before the start of neurogenesis (Crotti and Ransohoff, 2016). Once colonization has occurred, microglia proliferate in the brain until p14, after which numbers decrease until reaching stable adult levels (Nikodemova et al., 2015).

Microglia share the same spatiotemporal compartment as NECs, and are instrumental for proper brain development. One of the main roles of microglia is to contribute to the fine-tuning of the synapsis through a process called pruning, and also phagocytize axons and dying or dead cells (Schafer and Stevens, 2013). Towards the end of neurogenesis, microglia also remove neural progenitors in order to control the size of the progenitor pool (Cunningham et al., 2013). In the adult brain, microglia are constantly surveilling the microenvironment (Napoli and Neumann, 2009), and similarly to macrophages

can acquire a pro-inflammatory phenotype that if prolonged in time or in the absence of an insult can lead neurotoxicity (Dheen et al., 2007).

Neurodevelopmental disorders

Neurodevelopmental disorders (NDDs) are an heterogenous class of disorders with disruptions in the tightly regulated processes during brain development and function. NDDs are characterized by a vast genetic and clinical variability and affect >3% of children worldwide. NDDs have a very heterogenous etiology and are characterized by deficits in cognition, communication, adaptive behavior and psychomotor development. NDDs include autism spectrum disorder (ASD), intellectual disability (ID), and attention deficit hyperactivity disorder (ADHD), amongst others (Parenti et al., 2020). Although the etiology of many NDDs is unknown, many causes have been associated with increased risk of NDD development, including genetic, environmental, and infectious factors. The co-occurrence of several NDDs in the same patient have been frequently reported, suggesting the existence of common biological mechanisms (Berg and Geschwind, 2012, Singh et al., 2017a, Cardoso et al., 2019).

Genetic components and interactome of NDDs

The knowledge on the genetic components of NDDs has exploded in recent years thanks to the large-scale sequencing of genomes of both affected individuals and their healthy relatives. This allowed for the identification of all types of lesions including copy number variants (CNVs), indels, and disrupting point mutations, as well as the definition of common hereditary variants and *de novo* mutations (Sebat et al., 2007, Malhotra et al., 2011, Levy et al., 2011, Sanders et al., 2012, Hamdan et al., 2014, Niemi et al., 2018). Most of these mutations are found in protein-coding genes that are preferentially expressed in the developing brain and are intolerant to haploinsufficiency. The majority of the identified NDD susceptibility genes are part of a few molecular pathways that include chromatin remodeling, synaptic plasticity and function, splicing and transcriptional regulation, and immunity (Voineagu et al., 2011, Parikshak et al., 2013, Iossifov et al., 2014, De Rubeis et al., 2014).

The relationship between NDD-susceptibility genes has been shown not to be restricted to the molecular pathways and gene regulatory networks. Interactomes built computationally and experimentally from NDD-susceptibility genes

showed that proteins encoded by NDD genes have high connectivity (Cristino et al., 2014, Pinto et al., 2014, Satterstrom et al., 2020) and this is enriched in protein isoforms that are expressed predominantly in the brain (Corominas et al., 2014). This suggest that a mutation or CNV in one NDD gene can technically also affect other NDD genes within these protein-protein interaction (PPI) networks. Furthermore, large-scale transcriptomic studies performed on postmortem brains from idiopathic ASD and other major psychiatric disorders (schizophrenia, depression, and bipolar disorder) identified an upregulation of immune genes and microglial markers. In ASD and schizophrenia there was also a downregulation of neuronal and synaptic modules (Voineagu et al., 2011, Gupta et al., 2014, Akbarian et al., 2015, Gandal et al., 2018a).

Inflammation in NDDs

In the last decades, a link between maternal immune activation (MIA) and the development of NDDs has been established. It started with the observation that children with congenital rubella had a higher than expected incidence of ASD, and that 20% of children born to mothers who had rubella during pregnancy developed schizophrenia as adults (Chess, 1971, Brown, 2006). Later, it was also observed that children born to mothers who suffered from different infections (such as influenza or cytomegalovirus) during pregnancy had an increased risk of developing several NDDs, leading to a seminal study that demonstrated a clear association between maternal viral infection during the first trimester and bacterial infection during the second trimester of gestation and ASD (Atladottir et al., 2010). These studies established that immune activation, rather than a certain insult, had the potential to affect neurodevelopment (Reviewed in Solek et al., 2018).

Post-mortem studies from brains of ASD individuals showed signs consistent with CNS inflammation and immune system dysregulation, including increased numbers and activation of microglia and astrocytes, and the upregulation of pro-inflammatory cytokines (Pardo et al., 2005, Morgan et al., 2010).

MIA can be mimicked in mouse models by injecting lipopolysaccharide (LPS) or the viral mimic poly(I:C) to pregnant females to generate an antibacterial or antiviral immune response, respectively, through interaction with the Toll-like receptors (TLRs). The male MIA offspring presents with deficient sociability and communication, as well as increased repetitive behaviors, which are classical autism hallmarks (Malkova et al., 2012). These abnormal behaviors and changes

in gene expression are dependent on the proinflammatory cytokine interleukin-6 (IL-6), whose expression is induced by MIA. Blocking IL-6 or elevation if the anti-inflammatory cytokine IL-10 can reverse these effects (Estes and McAllister, 2016). MIA-induced by LPS injection causes an increase in the number of microglia in the developing neocortex of the offspring. This increase in microglia has been associated with a decrease in NECs, as microglia controls the number of neural progenitors by phagocytosis (Reviewed in Tong and Vidyadaran, 2016).

Microexon splicing defects in NDDs

Alternative splicing (AS) is a regulatory mechanism by which different exons from a gene are incorporated to produce multiple transcripts, therefore expanding the transcriptome and proteome (Braunschweig et al., 2013). In humans, 95% of multi-exon genes undergo AS, resulting in splice variants that are variably expressed across cell and tissue types (Wang et al., 2008). AS is especially important in the nervous system, where it is tightly regulated and generates a vast transcriptomic complexity (Kim, Magen and Ast, 2007; Irimia and Blencowe, 2012).

AS takes place through a large protein-RNA machinery, the spliceosome, that differentially selects alternative exons, 5' and 3' splice sites, and introns. The assembly of spliceosomes at 5' and 3' splice sites is regulated by RNA-binding proteins (RBPs). Mutations in RBPs have been linked as a cause or contributor to several diseases, and so has the misregulation of AS (Scheckel and Darnell, 2015).

The main component of misregulated AS in neurological disorders are 3 to 27 nucleotide (nt) microexons (Quesnel-Vallières et al., 2015). Microexons are highly conserved through evolution, and usually frame-preserving. Although microexons represent only 1% of all AS events, they account for a third of all neural-related AS (Irimia et al., 2014).

Microexons are enriched in protein domains that are surface accessible, suggesting that they can regulate PPIs (Irimia et al., 2014; Li et al., 2015). Microexons are significantly enriched in genes that have been genetically linked to ASD, learning difficulties, and ID, and have been shown to be frequently misregulated in the brains of autistic individuals - up to 30% of known microexons are skipped in ASD (Irimia et al., 2014). Most neural microexons are regulated by the neuronal-specific Ser/Arg repeat-related protein of 100 kDa (nSR100/SRRM4), whose levels are reduced and associated with the microexon skipping seen in ASD patients (Irimia et al., 2014; Quesnel-Vallières et al., 2015).

Mice haploinsufficient for nSR100/SRRM4 display hallmarks of ASD, altered neuronal excitability and synaptic transmission. nSR100 is regulated by neuronal activity, with neuronal activation producing a reduction in nSR100 levels and splicing alteration (Quesnel-Vallières et al., 2016).

The Tousled-like kinases (TLKs)

Discovery of the Tousled kinase and the TLKs

The Tousled (*TSL)* gene was first identified in 1993 in *Arabidopsis thaliana* by mutational analysis as a gene required for the correct leaf and flower development (Roe et al., 1993). Loss of TSL resulted in disorganized flower and leaf development, hence the name Tousled. TSL homologues were subsequently identified in *C. elegans*, *T. brucei*, *D. melanogaster* and humans (Silljé et al., 1999, Carrera et al., 2003, Han et al., 2003, Li et al., 2007). These Tousled-like kinases (TLKs) are constitutively expressed in most cells and organs from plants to animals, suggesting a crucial role in development. Indeed, the single TLK genes in *C. elegans* (TLK-1) and in *D. melanogaster* (tlk) are essential for viability (Carrera et al., 2003, Han et al., 2005). In *C.elegans*, TLK-1 was found to have a critical role in transcription during early development (Han et al., 2003), and a role in chromosome segregation during mitosis, through interacting with Aurora B kinase (Han et al., 2005). In *D. melanogaster*, the loss of tlk resulted in abnormal nuclear divisions leading to cellular apoptosis (Carrera et al., 2003). Tlk was found to control cell cycle progression through the regulation of chromatin dynamics via its interaction and phosphorylation of the histone H3/H4 chaperone ASF1 (Carrera et al., 2003), as did yeast TLK-1 (Li et al., 2001). Mammals have two TLK genes (TLK1 and TLK2) that are located in different chromosomes (in humans 2q31.1 and 17q23.2, respectively) and encode for several isoforms of unknown relevance. An additional translationally regulated form of TLK1 that lacks the first 169 aa, named TLK1B, has been characterized (Sunavala-Dossabhoy et al., 2004). Consistent with data from lower organisms, mammalian TLKs show maximal expression and activity during S phase, coinciding with DNA replication (Silljé et al., 1999), and are suggested to have roles in DNA repair and organism development (Segura-Bayona et al., 2017, Sunavala-Dossabhoy, 2018).

TLKs domains, structure and activation mechanism

Sequence alignment performed in 1993 showed that TSL falls into a distinct family of kinases (Roe et al., 1993). TLKs are conserved through evolution from *Arabidopsis thaliana* to *Homo sapiens*, although they are absent in *Saccharomyces cerevisiae* (Silljé et al., 1999). In humans, TLKs represent a unique kinase family, located in the kinome between the Polo and AGC kinase families (Manning et al., 2002).

TLK proteins contain several domains, that are shared between the protein homologs. TLK structure is mainly divided into a C-terminus domain containing the Serine-Threonine protein kinase catalytic domain, and a large N-terminus regulatory domain (Silljé et al., 1999). Already in the 1990's it was suggested that *A.thaliana* TSL's N-terminus contains multiple coiled-coil structures (Roe et al., 1993) and a nuclear localization signal (NLS) (Roe et al., 1997) (Figure I3). Indeed, human TLK1 and TLK2 contain a predicted NLS in the N-terminus, a middle region containing three predicted coiled-coil (CC) domains, and a C-terminus kinase domain that shows 94% identity and 96% similarity between both kinases (Mortuza et al., 2018) (Figure I3). The C-terminus regions of the kinase domain (C-tail) contains many phosphorylation sites, one of which has been reported to inhibit TLK1 activity in response to DSBs following its modification by CHK1 (Groth et al., 2003, Krause et al., 2003). The last residues of the C-tail are predicted to be unstructured (Mortuza et al., 2018).

Coiled-coil domains are known to mediate protein-protein interactions (Alanis-Lobato et al., 2017). The oligomerization of TSL through the CC domains was suggested based on indirect yeast-two-hybrid experiments and reported to mediate catalytic activity (Roe et al., 1997). We showed that TLK2 interacts with TLK1 through the first of these domains (CC1) (Mortuza et al., 2018), confirming the presumed heterodimerization of TLK1 and TLK2 by (Silljé et al., 1999) and adding an extra layer of regulation. The deletion of CC1 also impairs the interaction with LC8, but not with ASF1 (Mortuza et al., 2018), its main known substrate and interactor (Figure I3). Neither the deletion of CC2 nor CC3 domains had a clear influence on these protein-protein interactions (Mortuza et al., 2018). TSL oligomerization has also been reported to be necessary for its catalytic activity (Roe et al., 1997). The deletion of individual CC domains does not have an impact of TLK2 activity on a generic substrate (MBP), suggesting that none of these domains, nor the heterodimerization with TLK1, are required for TLK2 activity. However, the physiological substrate, ASF1a, can only be phosphorylated by the TLK2 mutants that are capable of oligomerization,

suggesting that ASF1a recognition is dependent on the CC domains (Mortuza et al., 2018). It is also interesting to note that the deletion of the N-terminus results in increased autophosphorylation and substrate phosphorylation (Mortuza et al., 2018), suggesting it may play a negative regulatory role.

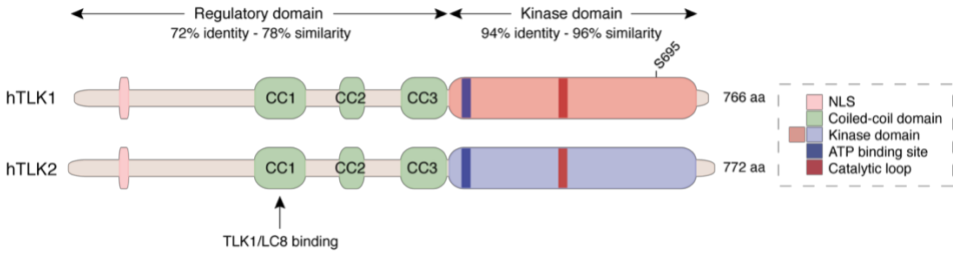


Figure I3. Domain architecture of human TLK1 and TLK2.

Human TLKs are highly homologous kinases composed of a C-terminus kinase domain and an N-terminus regulatory domain, which contains an NLS and three coiled-coil domains that are involved in protein-protein interactions. Adapted from (Segura-Bayona and Stracker, 2019).

The crystallization of the unphosphorylated TLK2 kinase domain in complex with ATPγS (Garrote et al., 2014) provided insights into the structural properties and mode of activation of TLK2. Key *cis* and *trans* autophosphorylation sites essential for its kinase activity were identified. A mechanism of activation that requires both *cis* and *trans* phosphorylation is proposed, suggesting that TLKs do not need an activating phosphorylation by an external kinase, in contrast to other kinases like CHK2 (Cai et al., 2009, Mortuza et al., 2018).

TLK2 is pre-activated in the monomer form through *cis*-autophosphorylation in the kinase domain. These events trigger a cascade of conformational changes that makes new sites available for further *cis*-autophosphorylation in the oligomerization domain and *trans*-phosphorylation also in the oligomerization domain and in the C-tail, in the context of a dimer conformation (Mortuza et al., 2018). These *cis*- and *trans*-phosphorylation events promote the assembly of higher order oligomers and lead to the full activation of the kinase (Mortuza et al., 2018) (Figure I4). The orientation of the oligomers' components remains unknown, but the fact that the CC domains are *trans*-phosphorylated suggest that the CC domains of a monomer could be located in close proximity to the kinase domain of another monomer (Mortuza et al., 2018).

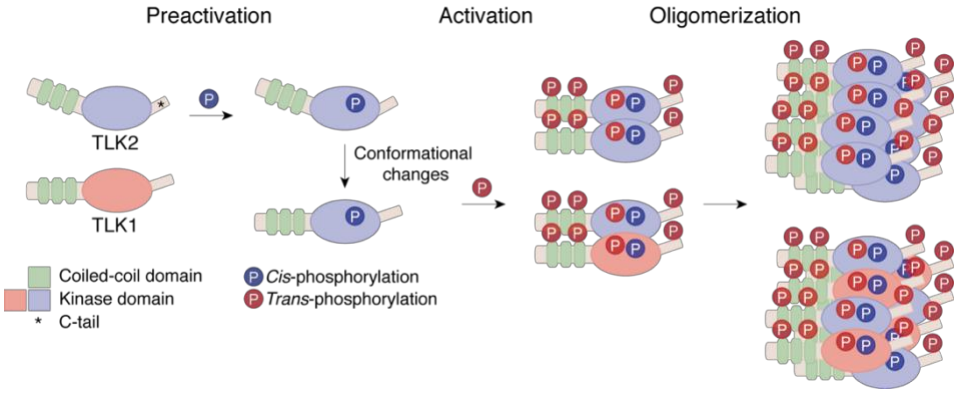


Figure I4. Model of human TLK activation.

A mechanism of activation that requires both *cis* and *trans* phosphorylation has been proposed based on structural and biochemical studies. Briefly, TLK2 is pre-activated in the monomer form through *cis*-autophosphorylation in the kinase domain, which trigger conformational changes that makes new sites available for further *cis*-autophosphorylation and *trans*-phosphorylation in the oligomerization domain and in the C-tail. These *cis*- and *trans*-phosphorylations promote the assembly of higher order oligomers, which are necessary for the full activation of the kinase. Adapted from (Mortuza et al., 2018).

Known substrates, interactors, regulators and functions of human TLKs

TLKs role in DNA replication, histone deposition, and chromatin maintenance

Human *TLK* expression is constitutive throughout the cell cycle at both mRNA and protein levels, but TLKs activity fluctuates and its activity peaks during S phase, coinciding with DNA replication (Silljé et al., 1999). Inhibition of DNA replication or DSBs during S phase transiently inhibits human TLKs activity in a DDR-dependent manner, and this reversible inhibition happens via direct phosphorylation of TLK1 at S695 by CHK1 (Figures I3 and I5A) (Groth et al., 2003; Krause et al., 2003).

The H3/H4 histone chaperone ASF1 has been identified as an interactor of TLKs in all organisms where it has been examined. TLKs have also been shown to phosphorylate ASF1 during replication both *in vitro* and *in vivo*, and TLK-phosphorylation sites in mammalian ASF1a and ASF1b have been mapped and functionally investigated. This TLK-mediated phosphorylation promotes ASF1's

binding affinity for H3-H4 heterodimers, which are then delivered to downstream histone chaperones for nucleosome assembly (Li et al., 2001, Silljé and Nigg, 2001, Carrera et al., 2003, Groth et al., 2005, Klimovskaia et al., 2014, Segura-Bayona et al., 2017, Hammond et al., 2017) (Figure I5A).

In cancer cells, acute depletion of TLK activity slowed replication fork progression and impaired *de novo* histone deposition of both H3.1 and H3.3 variants. This led to replication stress, characterized by replication fork stalling and ssDNA accumulation, with sustained TLK depletion leading to DSB, DDR activation and a p53-dependent G1 arrest (Figure I5B) (Lee et al., 2018). TLK depleted cells were hypersensitive to ATR, CHK1 or PARP inhibition, which strongly increased DNA damage and reduced cell viability, suggesting that the activity of proteins responsible to stabilize stalled replication forks is essential to avoid collapse in a TLK-deficient background (Lee et al., 2018).

The reduction in histone density upon TLK loss also results in chromatin decompaction and increased genome accessibility, especially at heterochromatic regions (Figure I5C). This is due in part to the reduction in H3.3-containing nucleosomes, as H3.3 is the variant required for the establishment of H3K9me3 marks. The desilencing of heterochromatin upon TLK loss resulted in the spurious transcription of repetitive elements, such as endogenous retroviruses (ERVs), satellites and telomeres, that triggered an antiviral transcriptional response that included the activation of Type 1 interferons (IFNs) and the tumor necrosis factor alpha (TNFα) pathway (Figure I5C)(Segura-Bayona et al., 2020). TLK loss also induces features of alternative-lengthening of telomeres (ALT), a recombination-based telomere maintenance mechanism used by 10%–15% of tumors (Hoang and O'Sullivan, 2020). TLK depletion causes classical ALT features, including increased DNA damage signaling at telomeres, an increase of ALT-associated PML bodies (APBs), where ALT-dependent telomere recombination takes place, and an increase in the production of extrachromosomal telomeric sequences (ECTRs). In addition, Stimulator of IFN gene (STING), a key regulator of interferon transcription, is epigenetically suppressed in many ALT+ cancers and was reactivated by TLK1/2 depletion. Together, TLK loss culminates in a Cyclic GMP–AMP synthase (cGAS)-STING-TANK binding kinase 1 (TBK1)-mediated innate immune response that is independent of RS, but attenuated by the depletion of factors required to produce

extratelomeric DNA, implicating ECTRs as the potential trigger (Figure I5C) (Segura-Bayona et al., 2020).

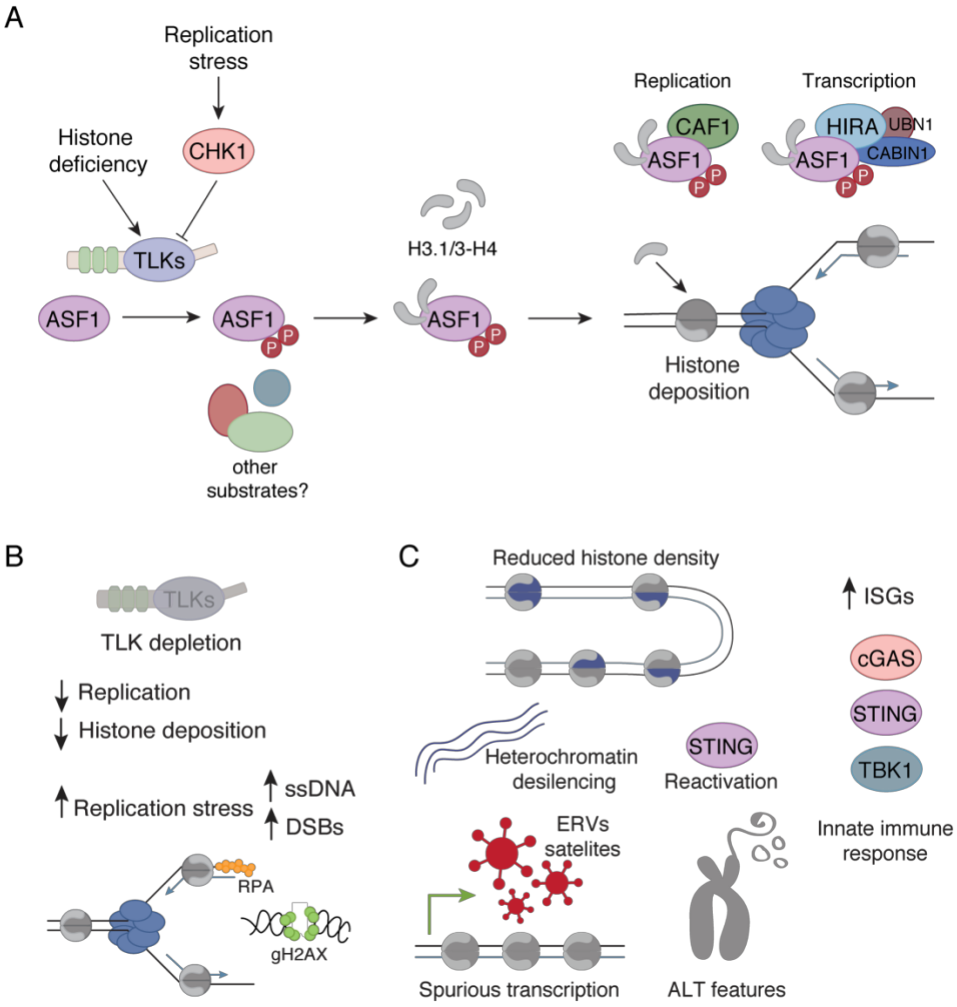


Figure I5. TLKs have a role in histone deposition and genome and epigenome stability.

(**A**) TLKs interact and phosphorylate the histone chaperones ASF1a and ASF1b, thus promoting ASF1's binding to H3/H4 heterodimers. ASF1 then delivers the histones to downstream histone chaperones such as CAF1 and HIRA, for replication-coupled and replication-independent chromatin assembly, respectively. TLK activity is triggered by histone deficiency and inhibited by the DDR. Adapted from (Klimovskaia et al., 2014). (**B**) Loss of TLK activity decreases replication and histone deposition, generating replication stress and DNA damage. (**C**) The reduction in histone density upon TLK loss

results in STING reactivation in cells where it is silenced and allows for a desilencing of heterochromatic regions, such as repetitive regions and telomeres, that leads to the spurious transcription or ERVs, satelites and other repetitive elements, as well as to an ALT phenotype characterized by extratelomeric DNA. The latter culminates in a cGAS-STING-TBK1-dependent innate immune response. Adapted from (Segura-Bayona et al., 2020). ERV, endogenous retrovirus; ALT, alternative lengthening of telomeres; ISG, interferon-stimulated genes.

The role of TLKs in DNA repair

Over 150 potential substrates of human TLKs have been proposed beyond ASF1 (Singh et al., 2017b), but only a handful of these have been clearly validated or functionally validated. Amongst the studied interactors are LC8/DYNLL1, NEK1, and RAD9.

We and others identified LC8-type 1 and 2 (DYNLL1 and DYNLL2) as consistent interactors of TLKs (Hein et al., 2015, Boldt et al., 2016, Segura-Bayona et al., 2017). LC8 was first identified as a subunit of the microtubule-based dynein motor complex (Rapali et al., 2011) but has since been shown to localize both in the nucleus and the cytoplasm and to interact with a large number of proteins unrelated to microtubule-based transport. LC8 has been proposed to be a multimerization hub essential to organize and stabilize many protein networks (Barbar, 2008), including complexes involved in DSB sensing such as the MRN complex, DNA resection such as BLM and DNA2, the DSB repair protein 53BP1, and other kinases, such as NEK9 (Lo et al., 2005, Regue et al., 2011, He et al., 2018, West et al., 2019). We have shown that the interaction between LC8 and TLK2 requires the CC1 domain of TLK2 (Mortuza et al., 2018) (Figure I3), and that LC8 is not phosphorylated *in vitro* by active TLK2 (Segura-Bayona et al., 2017), favoring the idea that LC8 is needed for the heterodimerization of TLK1 and TLK2.

TLK1 has been shown to phosphorylate RAD9, a component of the 9-1-1 clamp loader that is important for initiating and maintaining the checkpoint response. RAD9 regulates DNA damage-induced CHK1 activation via binding to TOPBP1 (Delacroix et al., 2007). TLK1 and RAD9 interact constitutively and TLK1 phosphorylates RAD9 at T355, phosphorylation that is reduced upon exposure to ionizing radiation. On the other hand, TLK1B has been shown to

phosphorylate RAD9 at S328, which promotes RAD9 dissociated from the chromatin and redistribution to the cytoplasm, causing DSB and further activation of the DDR (Sunavala-Dossabhoy and De Benedetti, 2009, Canfield et al., 2009, Kelly and Davey, 2013, Awate and De Benedetti, 2016).

TLK1B phosphorylates NEK1 on T141, which activates NEK1, and the interaction is enhanced following oxidative stress induced by exposure to H_2O_2. Overexpression of a NEK1 T141A mutant or TLK1 inhibition influenced cell cycle checkpoint regulation in response to oxidative stress. Recently, it has been shown that NEK1 phosphorylation by TLK1 also promotes mitochondrial integrity, and that treating NEK1-T141A mutant cells with doxorubicin increases apoptosis. This suggested that targeting the TLK1/NEK1 axis might be a novel therapy for advanced prostate cancer, which frequently express high levels of NEK1 (Singh et al., 2017b, Singh et al., 2019a, Singh et al., 2019b, Singh et al., 2020).

TLKs in mammalian development and disease

TLKs during murine development

The single copy of a TLK homolog in lower organisms is essential for proper development or viability (Roe et al., 1993, Carrera et al., 2003, Han et al., 2005). The generation of *Tlk1*- and *Tlk2*- deficient mice showed that, while TLK1 was dispensable for murine viability, mice lacking TLK2 were embryonic lethal due to placental failure. Histological analysis of *Tlk2*$^{-/-}$ placentas did not show aberrant proliferation or increased DSBs or apoptosis, but showed defects in trophoblast cell differentiation. *Tlk2*$^{-/-}$ placentas showed strongly reduced levels of ASF1 phosphorylation, as the low levels of TLK1 protein in the placenta were not able to support the necessary levels of ASF1 phosphorylation needed for proper trophoblast differentiation. The bypass of the placental phenotype using a conditional Tlk2 allele coupled to a Sox2-Cre system allowed for the generation of adult *Tlk2*$^{-/-}$ animals with no apparent overt phenotype except for a mild developmental delay (Segura-Bayona et al., 2017).

Across adult tissues, relative mRNA levels of *Tlk1* and *Tlk2* were similar, with the exception of the testis, where *Tlk2* levels were higher. TLK1 and TLK2 protein levels were generally consistent with the mRNA levels, suggesting that

TLK1 and TLK2 have redundant roles in adult tissues. This was supported by the lack of a clear phenotype in *Tlk1* or *Tlk2* deficient adult mice, in contrast to the placenta, where TLK2 is essential likely due to the low levels of TLK1 (Segura-Bayona et al., 2017).

TLKs in cancer

DNA amplification leading to protein overexpression is a common event in cancer. This is frequent for oncogenes such as HER2 (encoded by *ERBB2*), which is amplified and overexpressed in approximately 20% of breast cancers and linked to cancer progression and pathogenesis (King et al., 1985, Aulmann et al., 2006, Hayes, 2019).

In 1994, several chromosomal regions were described to have a high frequency of amplification in breast cancer, containing unknown genes at the time predicted to contribute to cancer progression (Kallioniemi et al., 1994). These regions included 17q23, where *TLK2* is located. It is now clear that 17q23 contains several oncogenes, including *TBX-2*, *RAD51C*, and *RPS6KB1* (Kelemen et al., 2009), and that at least one region of 17q23 is amplified in 43% of all breast tumors (Sinclair et al., 2003). In the case of *TLK2*, it is amplified in ~9% of all breast cancers, a percentage that goes up to 21.3% in the estrogen receptor positive (ER^+) luminal B subtype (Kim et al., 2016, Lee et al., 2018). *TLK2* amplification is more significant in advanced and aggressive tumors (Kim et al., 2016), and can also be observed in a number of tumor types beyond breast tumors (Figure I6A) (Lee et al., 2018). *TLK2* amplification does not correlate with that of most known oncogenes amplified in breast cancer (*MYC*, *CCND1*) or with *ERBB2*, which is located in 17q12 (Kim et al., 2016).

High expression levels of *TLK2* have been associated with poor patient outcome in subsets of breast cancer patients (Figure I6D) and other tumor types such as liver and cervical cancer (Figures I6B and I6C) (Kim et al., 2016, Lee et al., 2018), and TLK2 overexpression has been shown to increase cell migration and invasiveness *in vitro* (Lin et al., 2019, Kim et al., 2016), characteristics associated with metastasis. TLK2 was found to be highly phosphorylated in proteogenomics studies, suggesting that a part from being amplified it might also have increased activity (Mertins et al., 2016).

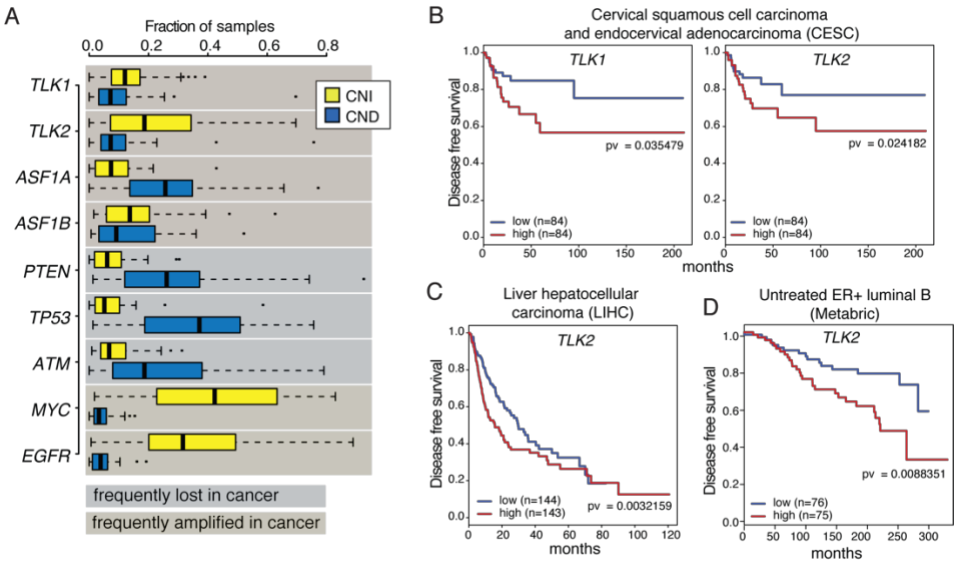


Figure I6. TLKs are amplified in cancer and in some cases high TLK2 expression correlate with reduced disease-free survival.

(**A**) Box plots of the fraction of patients with copy number increase (CNI) or copy number decrease (CND) among all TCGA cohorts for the indicated genes. The tumor suppressor genes PTEN, PT53 and ATM are shown as an example of genes with frequent CND in cancer, while the oncogenes MYC and EGFR frequently have CNI. (**B-D**) Kaplan-Meier plot of multivariate disease-free survival analysis of the TCGA cohorts cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) (**B**), liver hepatocellular carcinoma (LIHC) (**C**), and Untreated ER+ luminal B breast cancer (Metabric cohort) (**D**) based on expression of the indicated gene (*TLK1* or *TLK2*). Adapted from (Lee et al., 2018).

TLK inhibitors for cancer treatment

Protein kinases dysregulation has been demonstrated to be directly involved in the pathogenesis of many illnesses including cancer (Ferguson and Gray, 2018). For this reason, kinases have been in the center of drug discovery programs in the pharmaceutical industry for 30 years, with 52 small molecule protein kinase inhibitors approved to date by the Food and Drug Administration (FDA) (Roskoski, 2020). Most of these drugs are approved for the treatment of cancer, such as imatinib, a BCR-Abl tyrosine kinase inhibitor used to treat leukemias that are Chromosome-Philadelphia positive, amongst other cancer types (Druker et al., 2001, Ferguson and Gray, 2018).

Due to the implications in some cancers and the fact that they are kinases, TLKs have been suggested as promising druggable targets for therapeutic intervention in cancer (Kim et al., 2016, Lee et al., 2018, Lin et al., 2019). However, TLK-specific inhibitors have not yet been identified. Potential compounds have been found after screening several available compound libraries (Ronald et al., 2013) and publicly available kinase profiling data sets (Gao et al., 2013). The first suggested drugs to inhibit TLK1 *in vitro* belong to the family of antipsychotics phenothiazine, including Thioridazine hydrochloride (THD), Perphenazine (PPH), and Promazine hydrochloride (PMZ). They are ATP-competitors and are known to bind many kinases with high affinity and low selectivity. They were reported to increase γH2AX in cancer cells and slightly reduce tumor growth in mice treated in combination with the chemotherapeutic agent doxorubicin (Ronald et al., 2013).

Small-molecule inhibitors predicted to have inhibitory activity against TLK2 have been identified from publicly available kinase profiling data sets (Gao et al., 2013). Molecules in this screen were tested in 2 studies: two protein kinase C (PKC) inhibitors (Gö 6983 and GF 109203X) (Kim et al., 2016), CDK1 inhibitor (CGP74514A), GSK-3 inhibitor (Inhibitor XIII), and the Indirubin derivative E804 (Mortuza et al., 2018) were shown to inhibit TLK2 *in vitro* although at high concentrations, making them unfit for *in vivo* studies. Another drawback of these molecules is that they are highly promiscuous, thus increasing the need of target selectivity in order to reduce off-target effects (Cohen and Alessi, 2013).

In 2018, the crystal structure of TLK2's kinase domain was published and used to model small-molecule inhibitors using HADDOCK docking (van Zundert et al., 2016, Mortuza et al., 2018). This allowed for the identification of the residues through which the inhibitors interact with the kinase. These preliminary results showed that the kinase domain crystal structure can serve as a platform for a rational design and successful screening of TLK2 inhibitors (Figure I7) (Mortuza et al., 2018).

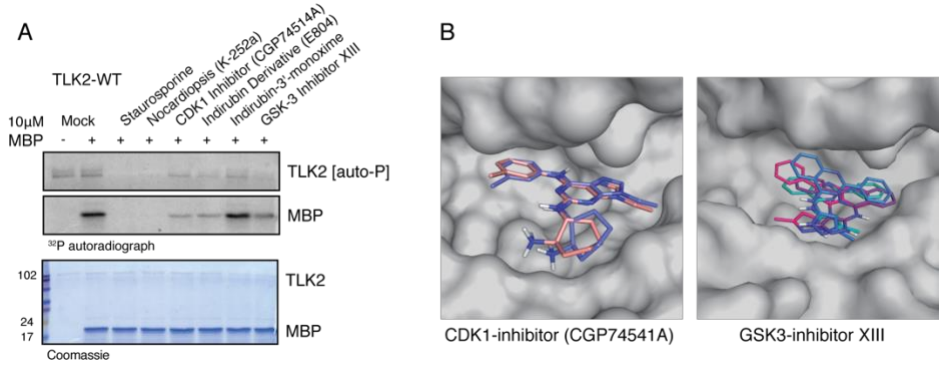


Figure I7. Testing of TLK2 inhibitors.

(**A**) *In vitro* kinase assay of TLK2 pulled down from AD293 cells using myelin basic protein (MBP) as a substrate and the indicated kinase inhibitors at 10 μM. Coomassie staining shows equal loading. (**B**) Detailed view of the ATP-binding pocket in the TLK2 kinase domain crystal structure showing the modelled molecules that inhibit TLK2 kinase activity in (**A**). Adapted from (Mortuza et al., 2018).

TLKs in neurodevelopmental disorders

Genetic factors such as inherited and *de novo* mutations (DNM) affecting protein-coding genes are thought to be the most significant cause of neurodevelopmental disorders (NDD), such as ID and ASD. Recent studies involving genome-wide interrogation have led to the identification of a large and fast-growing number of NDD candidate genes (Abrahams et al., 2013, Iossifov et al., 2014, Lelieveld et al., 2016, Schluth-Bolard et al., 2019, Satterstrom et al., 2020). Amongst those, *TLK2* was identified as one of ten new candidate genes in a meta-analysis of more than 2,600 *de novo* mutations (DNMs) from over 2,000 patients (Lelieveld et al., 2016). DNMs in *TLK2* had already been identified previously in spontaneous ASD across different ethnicities by sequencing parent-child trios (O'Roak et al., 2011, Takata et al., 2018). (Figure I8, Supplementary table 2). In 2018, a very complete study identified *de novo* and inherited heterozygous mutations in *TLK2* in 38 unrelated individuals with 2 affected mothers from 26 different centers in 7 countries. These mutations in *TLK2* caused a distinct NDD, characterized by clinical features such as mild neurodevelopmental delay, ID, ASD, behavioral disorders, facial dysmorphism, skeletal malformations and severe gastrointestinal problems, named Mental Retardation Autosomal Dominant (MRD57, MIM: 618050) (Reijnders et al., 2018). A broad spectrum of variant types was identified, including 10 nonsense variants, 9 missense variants located mostly at known functional domains (CC domains and catalytic domain), 12 canonical splice-site

variants, 4 frameshift variants, and one balanced translocation resulting in a breakpoint between *TLK2* exons 2 and 3. Two of the missense mutations were recurrent amongst unrelated patients. All of the known *TLK2* DNMs are in heterozygosis, with the exception of an homozygous *TLK2* mutation described recently (Töpf *et al.*, 2020). Most of the mutations resulted in haploinsufficiency and were predicted to be loss-of-function (LOF) (Reijnders et al., 2018). A few of these mutations have been tested *in vitro* and resulted in a loss of TLK2 kinase activity (Mortuza et al., 2018).

Furthermore, TLK2 was recently identified in a genome-wide CRISPR-based screen as one of ~200 novel regulators of neuronal microexon splicing. TLK2 displays a modest effect as a positive regulator of nSR100/Srrm4 gene expression (Gonatopoulos-Pournatzis et al., 2018), whose levels are reduced and associated with the microexon skipping seen in ASD patients (Irimia *et al.*, 2014; Quesnel-Vallières *et al.*, 2015).

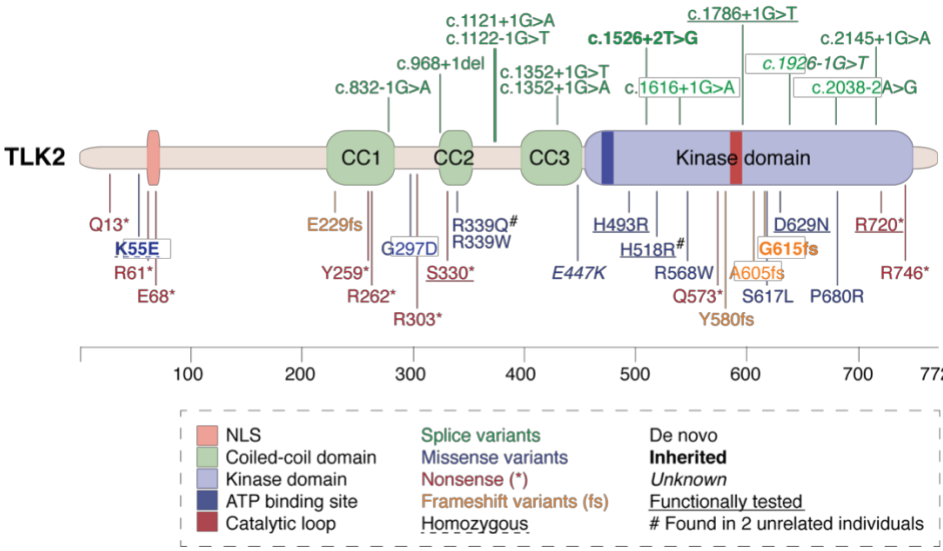


Figure I8. TLK2 mutations in patients with NDDs.

Diagram of the human TLK2 protein (Q86UE8-1; isoform 1) mapping all the identified mutations in patients with NDD in the following studies (O'Roak et al., 2011, Iossifov et al., 2014, Lelieveld et al., 2016, Reijnders et al., 2018, Satterstrom et al., 2020, Topf et al., 2020). The underlined mutations have been functionally tested in (Mortuza et al., 2018) or (Reijnders et al., 2018). Figure adapted and updated from (Segura-Bayona and Stracker, 2019).

Aims

Tousled-like kinases (TLKs) are implicated in organismal development and play a role in the pathogenesis of cancer and a distinct neurodevelopmental disorder (MRD57). We hypothesized that TLKs are a potential druggable target in cancer and that mutations in *TLK2* might contribute to MRD57 by causing chromatin abnormalities, a misregulation in microexon splicing and/or by eliciting an inflammatory response.

The specific aims of this thesis are:

- To determine the effect of *Tlk2* depletion in breast cancer progression and metastasis *in vivo*.

- To explore the potential of TLK inhibition as a therapeutic intervention for cancer treatment.

- To determine the role of TLK2 during mouse development and investigate how *Tlk2* loss contribute to the pathogenesis of neurodevelopmental disorders.

- To characterize the pathomechanism of TLK2 mutations identified in MRD57 patients.

- To establish the proximal interactome of TLK2 and how it is affected by MRD57 mutations.

Materials and methods

Animal studies

TLK2$^{F/F}$ Sox2-Cre mice were generated as described previously (Segura-Bayona et al., 2017) in the IRB Barcelona Mutant Mouse Core Facility. MMTV-PyMT-expressing TLK2$^{F/F}$ UbiquitinC-Cre^{ERT2} (UbCre^{ERT2}) mice were generated by crossing TLK2$^{F/F}$ mice with MMTV-PyMT (Guy et al., 1992, Attalla et al., 2020) and UbCre^{ERT2} mice (Gruber et al., 2007). MMTV-PyMT mice were generated by William Muller (McGill University) and obtained from Angel Nebreda (IRB Barcelona), while UbiquitinC-Cre^{ERT2} mice were obtained from the IRB Barcelona mouse mutant facility. All animals were maintained on a mixed 129/SvEv-C57BL/6 background strongly enriched for C57BL/6.

In order to obtain embryos from the desired embryonic day, timed matings were set up in the evening and plugs checked in the morning. Pregnant female mice were euthanized with CO_2 and embryos by decapitation.

For tumor studies, tumor size was analyzed by measuring with a digital caliper tumor length (L), width (W) and volume (V) were calculated on the basis of the following formula: $V = L*W^2/2$. Tumors were measured and mice were weighed 2-3 times a week. Experiments were started when tumors reached 150-200 mm^3, which was around 3-4 months of age. For *Tlk2* deletion, 4-Hydroxytamoxifen (4OHT) (Sigma, H6278) was resuspended in 10% ethanol/90% corn oil and mice were treated for 5 consecutive day via intraperitoneal injection (0.1 mg 4OHT/g of animal/day). Mice were euthanized with CO_2 3 or 15-days post-4-OHT treatment.

Mice were housed in individual ventilated cages (IVC) in the animal facilities of the Barcelona Science Park (PCB). Protocols (CEAA-PCB-15-000069) were approved by the Animal Care and Use committee of the PCB (CEEA-OH-PCB) in accordance with applicable legislation (Royal Decree 53/2013, of 1st February 2013, based on Directive European 2010/63/EU). All efforts were made to apply to minimize use and suffering.

Genomic DNA isolation

Genomic DNA was extracted from adult mice or embryo tip-tails or tumors by lysing and digesting with Proteinase K (0.4 mg/mL proteinase K in 10 mM Tris-HCl, 20 mM NaCl, 0.2% SDS, 0.5 mM EDTA) overnight at 56ºC. DNA was recovered by isopropanol precipitation, washed in 70% ethanol, air-dried and resuspended in 10 mM TE buffer to use in PCR reactions.

Genotyping PCR

Genotyping was performed by PCR reactions with Taq polymerase (Biotools). The primers used for the genotyping of the different alleles are summarized in Table 1, and the PCR conditions in Table 2. PCR products were run in an agarose gel of the appropriate % and visualized using SyBR Safe DNA gel Stain (Invitrogen).

Table 1. Primers used for mouse genotyping.

Mouse allele	Primer ID	Sequence 5'-3'	Fragment size	Inter-pretation
TLK2	TLK2-Fw	CTCTTAACCACTGAGCCATCTC	~1.3 kbp ~1 kbp ~450 bp	Cond allele Tlk2^{+} allele Tlk2^{-} allele
	TLK2-Rv	CAGGACAGGGAAAGGCTAAA		
MMTV -PyMT	MMTV-PYMT Fw	CAACCCGAGTTCTCCAACAGATAC	63 bp	MMTV-PyMT allele
	MMTV-PYMT Rv	AACACAAGGATTTCGTCTTCTTCCA		
Cre	Cre-Fw	GCGGTCTGGCAGTAAAAACTATC	~100 bp	Cre allele
	Cre-Rv	GTGAAACAGCATTGCTGTCACTT		
Sox2 Cre	Sox2-Cre Fw	CCAGTGCAGTGAAGCAAATC	207 bp 165 bp	WT allele Sox2Cre
	Cre WT Fw	CTTGTGTAGAGTGATGGCTTGA		
	Cre common	TAGTGCCCCATTTTTGAAGG		
X + Y Chr.	X + Y Fw	TGGATGGTGTGGCCAATG	335 bp 253 bp	Male Female
	X + Y Rv	CACCTGCACGTTGCCCTT		

Table 2. PCR program for *Tlk2*, *Cre*, *MMTV-PyMT*, *Sox2-Cre* alleles and Y chromosome.

TLK2, Cre, MMTV-PyMT			***Sox2-Cre***			**Y chromosome**		
T (ºC)	**Time**	**C**	**T (ºC)**	**Time**	**C**	**T (ºC)**	**Time**	**C**
95	3 min		95	5 min		94	4 min	
95	1 min		95	1 min		94	30 sec	
58	1 min	35-40	58	45 sec	35	57	30 sec	36
72	1:45 min		72	45 sec		72	25 sec	
72	7 min		72	5 min		72	3 min	
8	hold		8	hold		8	hold	

T, temperature; C, cycles.

Tumor digestion

Tumors were dissected and minced with a blade until there were no pieces bigger than 1 mm^3. Tumors were incubated with digestion buffer (DMEM F12 supplemented with 2% FBS, 0.2 mg/mL Collagenase IV (C5138, Sigma), 40 µg/mL DNAse I (10104159001, Sigma), and 0.2 mg/mL Dispase II (D4693, Sigma) in a water bath at 37ºC for 20 minutes cycles. During every cycle, tumors

were mechanically disaggregated every 5 minutes. After the third mechanical force was applied, the fragments are left to sediment for 5 minutes at 37ºC. The supernatant was collected in FACS buffer (PBS supplemented with 2% FBS, 1% BSA, and 2 mM EDTA disodium salt dihydrate) and fresh digestion buffer was added. The mechanical forces were applied in the following order: 2x vortex, 1x vigorous pipetting with a cut 1 mL tip, and 2x vigorous pipetting with a 1 mL tip. The digested tumors were filters through a 100 μm mesh (Falcon Corning brand) and spun down. The pellet was resuspended in Ammonium chloride buffer (150 mM ammonium chloride, 10 mM potassium bicarbonate, and 0.1 mM EDTA disodium salt dihydrate, pH 7.2) and incubated for 4 minutes at room temperature in order to lyse red blood cells. The cells were filtered again through a 100 μm mesh and spun down. The pellet was resuspended and the cells were counted using an automated cell counter.

Flow cytometry

The Fc receptors were blocked using 2.5 μg/mL of TruStain FcX™ anti-mouse CD16/32 antibody (Biolegend 101319) in FACS buffer for 20 minutes at 4ºC, followed by antibody staining for 20 minutes at 4ºC protected from the light. Antibodies used are summarized in Table 3. For sorting, DAPI was used as a viability dye.

Sorting and Picoprofiling

2.000 tumor cells ($EpCam^+$), immune cells ($CD45^+$), tumor associated fibroblasts ($CD140^+$) and endothelial cells ($CD31^+$) were sorted with the antibodies listed in Table 3 in a BD FACSAria II (BD Biosciences). For RNA isolation, cells were sorted into Eppendorf tubes containing 45 μL of 1X Lysis buffer (20 mM DTT, 0.5% SDS, 0.5 mg/mL proteinase K in 10 mM Tris HCl pH 7.4), then incubated for 15 minutes at 65ºC, snap frozen and stored at -80ºC. RNA was recovered using magnetic beads (RNAClean XP, Beckman Coulter). cDNA synthesis, library preparation and amplification were performed as described by (Gonzalez-Roca et al., 2010). Briefly, RNA was reverse transcribed and amplified using the Whole Transcriptome Amplification (WTA) method (WTA2, Sigma-Aldrich). cDNA amplification was monitored with SYBR Green on a real-time PCR detection system (CFX96, Bio-Rad). The amplification was stopped at 23 cycles, when fluorescence reached a plateau. Subsequently, cDNA was purified using a commercial kit (PureLink Quick PCR Purification Kit, Invitrogen), and

quantified using a microvolume spectrophotometer (Nanodrop ND-1000, Thermo Fisher Scientific).

Table 3. Antibodies used for flow cytometry and sorting.

Antigen	Fluorochrome	Clone	Reference and source	Dilution
CD3	FITC	17A2	Biolegend 100204	1:200
CD4	BV605	GK1.5	Biolegend 100451	1:200
CD8	BV786	53-6.7	Biolegend 100750	1:200
NK1.1	BV421/PB	PK136	Biolegend 108731	1:200
CD19	PE-Cy7	6D5	Biolegend 115520	1:200
PD-1	BV711	EH12.2H7	Biolegend 329927	1:200
CD69	APC	H1.2F3	Biolegend 104514	1:200
CD45	AF700	30-F11	Biolegend 103128	1:200
F4/80	BV605	BM8	Biolegend 123133	1:200
CD11b	PE-Cy7	M1/70	BD Bioscience 552850	1:200
CD11c	BV421/PB	N418	Biolegend 117322	1:200
Ly6C	APC	HK1.4	Biolegend 128016	1:200
Ly6G	BV786	1A8	Biolegend 127645	1:200
PD-L1	PE	10F.9G2	Biolegend 124308	1:200
MHC-II	FITC	AF6-120.01	BD Bioscience 553551	1:200
FVD	APC-Cy7		eBioscience 65-0865-18	1:5000
CD31	PE-Cy7		Abcam ab46733	1:600 (s)
CD140a	PE	APA5	BD Bioscience 562776	1:100 (s)
CD45	APC-eFluor780	30-F11	eBioscience 47-0451-82	1:500 (s)
CD326 (EpCam)	BV605	G8.8	BD Optibuild 740389	1:300 (s)

(s), sorting.

Microarrays

For EpCam$^+$ cells transcriptional profiling, 8 μg of cDNA were fragmented and labelled (Gene Chip Mapping 250K Nsp Assay Kit, Affymetrix) according to the manufacturer instructions. Array hybridization was performed using the GeneChip Hybridization, Wash, and Stain Kit (Applied Biosystems). Briefly, libraries were denatured at 99°C for 2 minutes prior to incubation into the Clariom S mouse Assay array (Applied Biosystems). Libraries were hybridized on the arrays for 16 hours at 45°C for 60 rpm at GeneChip Hybridization Oven 645 (Affymetrix, Thermo Fisher Scientific). Washing and Stain steps were performed using a GeneChip Fluidics Station 450 following the instruments Clariom S Mouse Array protocol (Affymetrix/ThermoFisher Scientific). Finally, the arrays were scanned with a GeneChip Scanner GCS3000 (Affymetrix, Thermo Fisher

Scientific) using default parameters. The CEL files containing the microarray data were generated with the Command Console software (Affymetrix, Thermo Fisher Scientific), and were used for downstream bioinformatics analysis.

Microarray samples were processed using package oligo (Carvalho and Irizarry, 2010) from R (R Core Team, 2017) and Bioconductor (Gentleman et al., 2004). Raw cel files were normalized using RMA background correction and summarization (Irizarry et al., 2003) at the core transcript level. Chip probesets were annotated using the information provided by Affymetrix (Thermo-Fisher Scientific). Expression intensities were summarized at the gene level (gene symbol) by the scores from the first principal component of the probesets mapping to the same gene. The sign of theses score was corrected sot that it was congruent to the sign of the probeset contributing the most to the first component. This component was centered and scaled to the weighted mean of the probesets' means and standard deviations, where the contributions to this first component were used as weights. Standard quality controls were performed in order to identify abnormal samples and relevant sources of technical variability (Gentleman, 2005) regarding: a) spatial artifacts in the hybridization process (scan images and pseudo-images from probe level models); b) intensity dependences of differences between chips (MvA plots); c) RNA quality (RNA digest plot); d) global intentisity levels (boxplot of perfect match log-intensity distributions before and after normalization and RLE plots); and e) anomalous intensity profile compared to the rest of samples (NUSE plots, Principal Component Analysis). No samples were excluded according to the results of these quality control checks. Technical metrics described in (Eklund and Szallasi, 2008) were computed for quality control and for evaluation as potential confounders factors. For visualization purposes, expression intensities were corrected a-priori by amplification batch and RMA.IQR metric using a linear model in which the condition of interest (Flox-NoCre / WT-cre / KO) was included as covariate.

Genes were evaluated for differential expression across groups using a moderated t-statistics by empirical Bayes shrinkage as implemented in the limma R package (Ritchie et al., 2015). Adjustment by multiple contrasts was performed using the Benjamini-Hochberg False Discover Rate (FDR) method (Benjamini Y., 1995). Additionally, a GaGa hierarchical model (Rossell, 2009, Rossell, 2013) was conducted to discriminate genes with different profile expression across groups at 5% FDR. For the latter, expression data corrected a-priori by technical sources of variation (amplication batch and RMA.IQR) was used (see above). Gene expression profiles were centered, scaled and graphically displayed in a heatmap

using a blue-gray-red colour gradation, in which blue represented low expression, red indicated high expression and a high colour intensity corresponded to a higher extreme expression in absolute value. For clarity, the most extreme values were truncated to -1.5 and 1.5 in this graphic.

Tumor growth statistical analysis

Tumor growth rates were compared between genotypes using a linear mixed effects model, R package (nlme) (Box, 1994, Pinheiro, 2000) with an autoregressive (AR-1) block diagonal (indicating independence between subjects) covariance structure (Zavrakidis et al., 2020). The logarithm of the tumor volumes was considered as response variable, the time and the interaction between time and genotype were set as fixed effects, and the animal subject was set as random effect. Adjustment for multiple testing (single-step correction method) was performed using the R package multcomp (Hothorn et al., 2008).

Analysis of TCGA signatures

TCGA RNA-Seq datasets were downloaded from the legacy archive of the NCI GDC commons database (Grossman et al., 2016) and processed separately for each cancer type. Expression measures were expressed in RSEM (Li and Dewey, 2011) in this TCGA version, which were log2-transformed and quantile normalized. For patients with multiple instances, a single sample chosen at random was kept while the rest were excluded from further analyses. For a number of cancer types there were two different platforms available: Illumina HiSeq 2000 (HiSeq) and Illumina Genome Analyzer (GA); in such cases, duplicated samples across platforms were removed from the GA dataset while the HiSeq instance was kept in the final dataset. Samples whose expression values showed an unusual distribution compared to the rest of samples in their datasets were also excluded (ACC: TCGA-OR-A5L9-01A; HNSC: TCGA-D6-A6ES-01A and TCGA-CV-A45Q-01A; LAML: TCGA-AB-2955-03A, TCGA-AB-2986-03A, TCGA-AB-2816-03A, TCGA-AB-2955-03A, TCGA-AB-2986-03A and TCGA-AB-2816-03A; LIHC: TCGA-DD-A3A6-11A and TCGA-FV-A4ZP-01A; SKCM: TCGA-D3-A2JK-06A; UCEC: TCGA-BS-A0V4-01A; UVM: TCGA-WC-A885-01A; COAD: TCGA-A6-2679-01A and TCGA-AA-A004-01A). For each cancer type, expression matrices were corrected a-priori by platform, source center, and plate id, when suitable. For doing so, a linear model was fitted to the expression values gene wise in which platform and source center

were included as fixed effects. Regarding the sample's plate id, it was included in the models as a fixed or a random effect depending on the number of levels and sample size available for the cancer type under consideration. In some cases, additional variables were also included in the models in order to preserve signal with biological or clinical relevance (BLCA: diagnosis subtype; BRCA: PAM50 subtype; COAD and STAD: micro satellite instability status; KIRP: tumor type). Aneuploidy score and Stromal and Leukocyte fraction estimates were obtained from (Taylor et al., 2018). To estimate chromosomal instability (CIN) we computed signatures with the gene set CIN25 from (Carter et al., 2006). Gene signature scores were computed as the mean of all genes in the signature after scaling the expression matrix gene wise. The global signature was defined as the score associated with the gene signature containing all genes in the expression matrix. Correlation coefficients and p values were computed using the 'pcor.test' function from the ppcor R package (Kim, 2015). In order to avoid spurious correlations due to technical and global effects (Caballe´ Mestres, 2018), partial correlations were computed whenever a gene signature was involved with the global signature as adjusting variable. The scores associated to aneuploidy, stromal fraction and leukocyte fraction were transformed using the square root to ensure normality of the data for the linear model. P values were adjusted for multiple comparisons using Benjamini-Hochberg.

Tissue dissection for RNA-Seq

Placentas and whole brains of E13.5 or E18.5 embryos were harvested after decapitation and the telencephalons were dissected out under a binocular microscope and immediately snap frozen in liquid nitrogen. Tail-tips were taken for genotyping. Tissues were stored at -80ºC until needed.

RNA extraction from mouse tissues for RNA-Seq

Total RNA was extracted using the Maxwell simplyRNA purification kit (Promega) following the manufacturer's instructions. Briefly, tissues were thawed on ice, and placed in homogenization solution containing 1-Thioglycerol. Reagent DX (Qiagen) was added to avoid the formation of foam. Tissues were then disrupted by zirconium beads in a mechanical tissue disruptor (Precellys 24, Bertin Technologies). Lysis buffer was added, the sample was vortexed and the homogenate was placed on the Maxwell 16 LEV prefilled Cartridge (MCE). 10 µL of DNase I was added in the corresponding cartridge well. RNA was extracted

in an automated Maxwell 16 Instrument (AS2000) (Promega) and eluted in 40 µL of Nuclease-Free Water. RNA samples were stored at -80ºC.

RNA-Seq

RNA samples were quantified using a Nanodrop 8000 (Thermo Scientific). RNA integrity was measured with a 2100 Bioanalyzer Instrument (Agilent) using an RNA 6000 Nano Chip (Agilent). Only RNAs with a RIN (RNA Integrity Number) values > 9.20 were used. RNA-Seq libraries were prepared using NEBNext® Poly(A) mRNA Magnetic Isolation Module (E7490, NEB) and NEBNext® Ultra™ II RNA Library Prep Kit (E7770, NEB) according to the manufacturer's instructions. In brief, ribosomal RNAs (rRNAs) were efficiently removed using oligodT-based magnetic beads, mRNA isolated and fragmented, and cDNA generated. Each of the libraries was labelled with a specific barcode and amplified by PCR in order to obtain an optimal yield (8 amplification cycles for the telencephalons and 9 for the placentas). Libraries were quantified and size measured with a 2100 Bioanalyzer Instrument (Agilent) using a DNA High Sensitivity Chip (Agilent). Equimolar pools were generated with the libraries and the pools were sequenced in a HiSeq 2500 sequencing system (Illumina) at the CNAG Sequencing Unit (Barcelona, Spain). For the telencephalons, 18 libraries were pooled and sequenced in 4 lanes 125 bp paired-end with an average of 70 million reads per sample, while for the placentas 9 libraries were pooled and sequenced in 1 lane 50 bp single-end with an average of 30 million reads per sample. Details regarding the samples used for RNA-seq are summarized in Table 4. Raw data was deposited in GEO (accession numbers GSE158424 and GSE158687).

Table 4. RNA-Seq sample information.

Tissue	Sample ID	Barcode number	Barcode seq	genotype	Sex	Litter
Telencephalon	1411	Index 1	ATCACG	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Female	Cx 1
Telencephalon	1412	Index 2	CGATGT	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Female	Cx 2
Telencephalon	1413	Index 3	TTAGGC	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Female	Cx 2
Telencephalon	1414	Index 4	TGACCA	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Male	Cx 2
Telencephalon	1415	Index 5	ACAGTG	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Male	Cx 2
Telencephalon	1416	Index 6	GCCAAT	*Tlk2*$^{-/-}$ Sox2Cre^{+}	Male	Cx 3
Telencephalon	1417	Index 7	CAGATC	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Female	Cx 3
Telencephalon	1418	Index 8	ACTTGA	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Female	Cx 3
Telencephalon	1419	Index 9	GATCAG	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Female	Cx 4
Telencephalon	1420	Index 10	TAGCTT	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Male	Cx 3

Telencephalon	1421	Index 11	GGCTAC	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Male	Cx 5
Telencephalon	1422	Index 12	CTTGTA	*Tlk2*$^{+/-}$ Sox2Cre^{+}	Male	Cx 4
Telencephalon	1423	Index 13	AGTCAA	*Tlk2*$^{+/+}$Sox2Cre^{+}	Female	Cx 6
Telencephalon	1424	Index 16	CCGTCC	*Tlk2*$^{+/+}$Sox2Cre^{+}	Female	Cx 5
Telencephalon	1425	Index 18	GTCCGC	*Tlk2*$^{+/+}$Sox2Cre^{+}	Female	Cx 7
Telencephalon	1426	Index 21	GTTTCG	*Tlk2*$^{+/+}$Sox2Cre^{+}	Male	Cx 7
Telencephalon	1427	Index 23	GAGTGG	*Tlk2*$^{+/+}$Sox2Cre^{+}	Male	Cx 8
Telencephalon	1428	Index 25	ACTGAT	*Tlk2*$^{+/+}$Sox2Cre^{+}	Male	Cx 8
Placenta	1541	Index 1	ATCACG	*Tlk2*$^{-/-}$	Male	Cx 9
Placenta	1542	Index 2	CGATGT	*Tlk2*$^{-/-}$	Male	Cx 9
Placenta	1543	Index 3	TTAGGC	*Tlk2*$^{-/-}$	Male	Cx 10
Placenta	1544	Index 4	TGACCA	*Tlk2*$^{+/-}$	Male	Cx 9
Placenta	1545	Index 7	CAGATC	*Tlk2*$^{+/-}$	Male	Cx 11
Placenta	1546	Index 8	ACTTGA	*Tlk2*$^{+/-}$	Male	Cx 12
Placenta	1547	Index 9	GATCAG	*Tlk2*$^{+/+}$	Male	Cx 11
Placenta	1548	Index 10	TAGCTT	*Tlk2*$^{+/+}$	Male	Cx 13
Placenta	1549	Index 11	GGCTAC	*Tlk2*$^{+/+}$	Male	Cx 13

Differential analysis and biological enrichment

Paired end reads were aligned to the mm10 version of the mouse genome using STAR (Dobin et al., 2013), version 2.3.0e, under default parameters. SAM files were sorted and indexed using Sambamba (Tarasov et al., 2015), version 0.5.9. The R (R Core Team, 2017) package Casper (Rossell et al., 2014), version 2.16.1, was used to quantify intensities at transcript and gene level. The GENCODE database was considered for transcript annotations. Differential expression between genotype condition pairs (KO vs WT, KO vs HET and HET vs WT) was performed using the Limma R package (Ritchie et al., 2015) on the gene level intensities. Genes were annotated to Hallmark terms (Liberzon et al., 2015), KEGG terms (Kanehisa and Goto, 2000) and GO terms (Mi et al., 2019) using the R package org.Hs.eg.db (Carlson, 2018). Gene set enrichment analysis was then performed employing the Casper gene level intensities, in which lowly expressed genes (genes with less than an average count of 5 reads) were filtered out. The rotation-based approach for enrichment (Di Wu, 2010) implemented in the R package limma was used to represent the null distribution. The maxmean enrichment statistic proposed in (Efron B, 2007), under restandardization, was considered for competitive testing. This same strategy for enrichment analysis was undertaken to test gene signatures defined by the Mouse Cell Atlas top 100 markers of every E18-brain cell type (Han et al., 2018), version 1.1.

Alternative splicing analysis

Vasttools (Tapial et al., 2017) version 2.2.2 was used for alignment and alternative splicing quantification. All observations from the same genotype were combined for PSI calculations. Only events with a minimum of 50 reads in both WT and KO samples were considered. We distinguished between exon skipping (from all Vasttools provided modules) in micro-exons (nt ≤ 27) and alternative-exons (nt > 27). Differential alternative splicing events were defined by a $|\Delta PSI| > 10$ in alternative-exons and by a $|\Delta PSI| > 5$ in micro-exons. Enrichment of lists of genes that at least had one differentially spliced event (separately for positive and negative effects) was tested against the SFARI database (Abrahams et al., 2013).

Repetitive element analysis

Annotations for repetitive elements were obtained from the repeatmasker website (Smit, 2013-2015). For every sample, counts per repetitive element were found using featureCounts (with arguments countMultiMappingReads=TRUE and fraction=TRUE), from Rsubread package (Liao et al., 2013). DESeq2 (Love et al., 2014) was considered to evaluate the differences between WT and KO samples.

Quantitative real-time PCR (qRT-PCR)

RNA was quantified using a Nanodrop 8000 Instrument (Thermo Fisher Scientific). The reverse transcription reaction was carried out using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) following the manufacturer's instructions, in a reaction volume of 20 µL and with random primers contained in the kit. cDNA was stored at −20°C. qPCR was performed using the comparative CT method and a Step-One-Plus Real-Time PCR Instrument (Applied Biosystems). SYBR Green reactions were carried out in technical triplicates in a final volume of 8 µL. For SYBR Green 1X SYBR Green PCR Master Mix (no. 4364344; Applied Biosystems), forward and reverse primer (Sigma-Aldrich; 100 nM each) and 30 ng of template were used. Thermocycling parameters used for SYBR Green reactions were: 95°C 20 seconds; 40 cycles 95°C 3 seconds, 60°C 30 seconds; melting curve. The 2−ddCT method was used for the analysis of the amplification products. mACTB was used as an endogenous control for normalization. Primer pairs used are indicated in Table 5. Data are presented as the mean±SEM of at least 2 experiments and 5 animals within each experimental group. Statistical significance was determined using an

unpaired t test with Welch's correction (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05).

Table 5. qRT-PCR primers used in this study.

Primer target	Sequence (5'-3')	Source
mActB_Fw	GGCTGTATTCCCCTCCATCG	(Terré et al., 2019)
mActB_Rv	CCAGTTGGTAACAATGCCATGT	
mH3f3b_Fw	CCAAGGCGGCTCGGAAAAGC	(Jang et al., 2015)
mH3f3b_Rv	GGTAACGACGGATCTCTCTCAGA	
mIfi27_Fw	TTCCCCCATTGGAGCCAAG	PrimerBank
mIfi27_Rv	AGGCTGCAATTCCTGAGGC	
mIfit2_Fw	CGGAAAGCAGAGGAAA TCAA	(White et al., 2016)
mIfit2_Rv	TGAAAGTTGCCATACCGAAG	
mSlc6a4_Fw	TATCCAATGGGTACTCCGCAG	PrimerBank
mSlc6a4_Rv	CCGTTCCCCTTGGTGAATCT	
mSRRM4_Fw	GGAAAGGGACCTTCAAAGCAG	PrimerBank
mSRRM4_Rv	GCTTGCGGGCTGTGATACT	

Protein extraction from mouse tissues for western blotting

Dissected tissues were thawed on ice and resuspended in TNG-150 buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Tween-20, 0.5% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). Tissues were disrupted by zirconium beads in a mechanical tissue disruptor (Precellys 24; Bertin Technologies) and protein concentration was quantified using the DC Protein Assay (Bio-Rad). Lysates were separated by SDS-PAGE and transferred to Nitrocellulose membranes (0.2 μm pore, Amersham Protran; Sigma-Aldrich). Membranes were blocked and antibodies prepared in 5% non-fat milk in PBS-T. The primary antibodies (Table 10) were detected with the appropriate secondary antibodies conjugated to Horseradish peroxidase (HRP) (Table 11) and visualized by ECL-Plus (GE Healthcare).

Histopathology

Adult lungs were inflated through the trachea with 10% neutral buffered formalin (Sigma-Aldrich) prior to harvest. Lungs and tumors were fixed overnight in 10% neutral buffered formalin at 4ºC and embedded in paraffin. Paraffin-embedded tissue sections (2-3 μm) were air dried and further dried at 60ºC overnight. In the case of the lungs, at least 20 serial cuts were done. 5 non-consecutive cuts were then dexawed and stained with hematoxylin and eosin (H&E) standard protocol

using a CoverStainer (Dako – Agilent) to be screened blindly for metastasis by a board-certified pathologist (Dr. Neus Prats).

For histopathology analysis of embryonic brains, we performed cardiac perfusion on E18.5 embryos with 4% paraformaldehyde (PFA) in phosphate buffer saline (PBS) pH 7.4. Embryonic heads were fixed in 4% PFA overnight at 4ºC, followed by cryoprotection in increasing concentration of sucrose in PBS (first 15%, then 30%), followed by an overnight incubation in a 1:1 solution of 30% sucrose and OCT (Tissue-Tek). Tissues were then embedded in OCT and frozen in liquid nitrogen-cooled isopentane. For tissue histological analysis of OCT samples, 10-12 μm thick cryosections were prepared and placed on glass slides.

Immunohistochemistry (IHC)

Immunohistochemistry for PyMT and CD45 was performed using a Ventana discovery XT for 60 minutes. Antigen retrieval was performed with Cell Conditioning 1 (CC1) buffer (Roche, 950-124) followed with the OmniMap anti-Rat HRP or with the Rabbit Anti-Rat for 30 minutes and the OmniMap anti-Rabbit HRP, respectively. Blocking was done with Casein (Roche, 760-219). Antigen–antibody complexes were reveled with ChromoMap DAB Kit (Roche, 760-159) or the Discovery Purple Kit (Roche, 760-229). Immunohistochemistry for Ki67 (60 minutes), H2A.X-pS139 (120 minutes), cleaved caspase 3 (120 minutes), and Iba-1 (120 minutes) were performed manually. For Ki67, sections were dewaxed as part of the antigen retrieval process using the low pH EnVision™ FLEX Target Retrieval Solutions (Dako – Agilent) for 20 minutes at 97ºC using a PT Link (Dako – Agilent), for H2A.X-pS139 and cleaved caspase 3 samples were dewaxed and antigen retrieval performed using citrate buffer pH6 for 20 minutes at 97ºC using a PT Link (Dako – Agilent). For Iba-1 samples were air dried, fixed with 10% neutral buffer formalin (Sigma-Aldrich) for 10 minutes at RT. Blocking was performed with Peroxidase-Blocking Solution at RT (Agilent, S2023) and 5% of goat normal serum (Life technology, 16210064) mixed with 2.5% BSA diluted in wash buffer for 10 and 60 minutes at RT. The secondary antibody used was the BrightVision poly HRP-Anti-Rabbit IgG, incubated for 45 minutes. Antigen–antibody complexes were reveled with 3-3'-diaminobenzidine (Agilent, K346811). Sections were counterstained with hematoxylin (Dako, CS700) and mounted with toluene-free mounting medium (Agilent, CS705) using a Dako CoverStainer. Specificity of staining was confirmed by staining with a Rat IgG or a Rabbit IgG isotype controls.

Information on the primary and secondary antibodies used can be found in Tables 10 and 11.

Brightfield images were acquired with a NanoZoomer-2.0 HT C9600 digital scanner (Hamamatsu) equipped with a 20X objective. All images were visualized with a gamma correction set at 1.8 in the image control panel of the NDP.view 2 U12388-01 software (Hamamatsu).

Cell lines and cell culture

AD-293 (Stratagene) and U-2-OS cells were grown in Dulbecco's modified Eagle's medium (DMEM; Thermo Fisher Scientific) supplemented with 10% (v/v) fetal bovine serum (FBS) (Sigma-Aldrich) and 50 U/mL penicillin/ 50 μg/mL streptomycin (Pen-Strep; Thermo Fisher Scientific). hTERT-immortalized human retinal epithelial cells (hTERT-RPE-1; ATCC) were grown in DMEM-F12 (Thermo Fisher Scientific) supplemented with 10% (v/v) FBS and Pen-Strep. Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood using Histopaque®-1077 (Sigma-Aldrich) and subsequently immortalized with Epstein-Barr Virus (EBV) and grown in Roswell Park Memorial Institute medium (RPMI-1640; Thermo Fisher Scientific) supplemented with 10% (v/v) FBS (Thermo Fisher Scientific), Pen-Strep and 1% (v/v) L-Glutamine. Cells were grown at 37ºC in a 5% CO_2 incubator. All tissue culture procedures were performed under sterile conditions in a Type II safety hood.

Small interfering RNA (siRNA)-mediated knockdown

hTERT-RPE-1 cells were transfected with 100 nM siRNAs (Sigma-Aldrich) using Lipofectamine RNAiMAX (Thermo Fisher Scientific) 24 hours after seeding. The transfection mix was prepared in Opti-MEM (Thermo Fisher Scientific) in the absence of antibiotics. Fresh media was added 8 hours post-transfection and cells were then cultured for 48 hours prior to harvesting. A siRNA targeting GFP was used as a mock control. The sequence of the siRNAs used are provided in Table 6.

Table 6. siRNA sequences.

siRNA target	Sequence (5'-3')	Source
siGFP	GGCUACGUCCAGGAGCGCCGCACC	(Rein et al., 2015)
siTLK1	GAAGCUCGGUCUAUUGUAA	(Lee et al., 2018)
siTLK2	GGAAAGGAUAAAUUCACAG	(Lee et al., 2018)

Genomic DNA extraction from human blood

Informed consent was obtained from participating families and the study protocol was approved by The University of Turin's internal Ethics Committee, according to the Declaration of Helsinki.

Genomic DNA (gDNA) was extracted from whole blood using the ReliaPrep Blood gDNA Miniprep kit (Promega) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientific). gDNA was amplified with touchdown PCR (annealing temperature 65-58°C) using KAPA2G Fast HotStart Taq (Merck). PCR products were further purified with FastAP Thermosensitive Alkaline Phosphatase following manufacturer's instructions.

Whole exome sequencing, array-based comparative genomic hybridization, prioritization, and variant calling

Whole Exome Sequencing (WES) data processing, variant filtering, and prioritization by allele frequency, predicted functional impact, and inheritance models were performed using an in-house implemented pipeline, which mainly takes advantage of the Genome Analysis Toolkit (GATK v.3.7). High-quality variants with an effect on the coding sequence or affecting splice site regions were filtered against public databases (dbSNP150 and GnomAD v.2.1) to retain: i) private and clinically associated variants; ii) annotated variants with an unknown frequency or having MAF <0.1%. The functional impact of variants was analyzed by Combined Annotation Dependent Depletion v.1.3 and using ACMG/AMP 2015 guidelines.

Array-based comparative genomic hybridization (array-CGH) was performed using a 60K whole-genome oligonucleotide microarray (Agilent Technologies). Families 1 and 2 were enrolled in the Autism Sequencing Consortium (ASC) project and their gDNA samples were sequenced at the Broad Institute on Illumina HiSeq sequencers as previously described (De Rubeis et al., 2014,

Satterstrom et al., 2020); variant calling was performed using targeted bioinformatic pipelines adapted for different pattern of inheritance.

Identified variants were confirmed by Sanger sequencing using standard conditions and the primers listed in Table 7. All variants are referred to GRCh37 annotation and to NM_001284333.2 (Supplementary table 2), in line with the previously published TLK2 structure (Mortuza et al., 2018). Variants have been deposited on ClinVar database with submission ID SUB8204586 (p.Asp551Gly), SUB8204660 (p.Glu475Ter), SUB8209291 (p.Ser617Leu) and SUB8204671 ((17)(q23.2)del).

Table 7. Primers used for Sanger sequencing of patient variants.

Variant	Primer	Sequence 5'-3'
c.1652A>G	Fw	AAAGGAGTGAGAAGCTGATGACC
	Rv	CAAAGAAAACCCACTAATACTGTCTCC
c.1423G>T	Fw	AGACCAGCACCCAAGTCC
	Rv	TGTATTCTGCTTGGTCACTTAGG

Fw (forward), Rv (reverse).

In silico prediction of missense variants impact

Variants were analysed with the VarSome tool (Kopanos et al., 2019) as a starting point for further analysis. The complete list of *in silico* predictors used can be found in (Pavinato et al., 2020). Variants frequencies were evaluated using Genome Aggregation Database (GnomAD) Browser version 2.1.1.

In silico modeling of TLK2 mutations

TLK2 mutations were modelled using the crystal structure of the TLK2 PKD (PDB: 5o0Y) (Mortuza et al., 2018) and the Structuropedia web interface for Modeller (Webb and Sali, 2016).

Single-cell gel electrophoresis (SCGE)

Samples were prepared according to the alkaline SCGE assay method as previously described (Flex et al., 2019). Briefly, lymphoblastoid cells (LCLs) harbouring the p.(Asp551Gly) variant and their wild-type counterpart were suspended in 0.7% low-melting agarose. Slides were prepared in duplicates and kept overnight in lysis solution at 4°C. Following lysis, slides were moved to

alkaline buffer for 20 minutes to unwind DNA. Electrophoresis was performed for 20 or 60 minutes at 20 V, 300 mA (0.8 V/cm) at 4°C. Slides were then neutralized in 0.4 M Tris pH 7.5 (3X 5 minutes), treated with absolute ethanol, stained with GelRed (Biotium), and analysed at a fluorescence microscope (Leica). To measure the extent of DNA damage induced by γ-ray irradiation and repair capability, LCLs were irradiated with 2-Gy γ-rays from a ^{137}Cs source (0.8 Gy/minute). During the treatment, cells were kept on ice to prevent DNA repair. Residual DNA damage was evaluated after 20 minutes at 37°C. Tail moments values were calculated by using a dedicated image analysis system (IAS2000 Delta System). For each experimental condition, at least 100 nucleoids were analysed. Statistical analysis was performed using two-tailed unpaired Student's t test.

Site-directed mutagenesis

TLK2 ASD mutants were generated using the QuickChange Lightning site-directed mutagenesis kit (Agilent Technologies) on the plasmids pcDNA3.1 N-SF-TAP-TLK2-WT (Mortuza et al., 2018) and BirA*-N-term-TLK2-WT (Segura-Bayona et al., 2017) following manufacturer's instructions. Primers used in the reaction are indicated in Table 8. Dpn I restriction enzyme (Agilent Technologies) was added to the PCR products to digest the parental supercoiled dsDNA. XL10-Gold ultracompetent cells were transformed with the resulting TLK2 ASDs' plasmids and colonies were selected with carbenicillin antibiotic. All constructs were sequenced (Macrogen) using the primers indicated in Table 9 and DNA was prepared with MaxiPreps (Promega).

Table 8. Site-directed mutagenesis primers.

Mutant	Primer	Sequence 5'-3'
TLK2-K55E c.163 A>G	Fw	AAGTAGAGACTCCCGAGGAAAAGCAGAATGACCAG
	Rv	CTGGTCATTCTGCTTTTCCTCGGGAGTCTCTACTT
TLK2-G297D c.887 G>A	Fw	GACCGCTTGAGACTGGACCACTTTACTACTGTC
	Rv	GACAGTAGTAAAGTGGTCCAGTCTCAAGCGGTC
TLK2-D551G c.1586 A>G	Fw	TTTCAGGTAGAAGTCCAGACCATTTCCCTCACAGTATTC
	Rv	GAATACTGTGAGGGAAATGGTCTGGACTTCTACCTGAA A
TLK2- S617L c.1784 C>T	Fw	CTATCATCATCCATGATCTTCAAAAGACCAAAATCTGTAA TTTTTATCTC
	Rv	GAGATAAAAATTACAGATTTTGGTCTTTTGAAGATCATG GATGATGATAG

Fw (forward), Rv (reverse).

Table 9. Primers used for Sanger sequencing.

Primer	Sequence 5'-3'	Sequenced plasmids
F1	CATGGAAGAATTGCATAGCCTG	TLK2-K55E
F480	AGAGCAGCTGGCGCAAAGGG	TLK2-G297D
F1526	CTTTTCACTGGATACTGAC	TLK2-D551G, TLK2-S617L

Transfection and Immunoprecipitation (IP) in mammalian cells

AD-293 cells were seeded in 15 cm plates and transiently transfected the next day with 20 μg of plasmid DNA using polyethylenimine (PEI) (Polysciences Inc., Warrington, PA) and 150mM NaCl. Medium was changed 6-8 hours post-transfection.

Cells were transfected with a series of pcDNA3.1 N-SF-TAP-TLK2 and BirA-N-term-TLK2 plasmids described above. In addition, a TLK2 kinase dead construct (KD, D592V) previously described (Mortuza et al., 2018) and pEXPR 105-Strep-ASF1a or -ASF1b (IBA GmbH), kindly provided by A. Groth, were used.

Cells were harvested 48 hours post-transfection and collected by scraping in PBS. Pelleted cells were lysed in 1 mL of ice-cold lysis buffer (50 mM Tris-HCl pH 7.5, 150mM NaCl, 1% Tween-20, 0.5% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)) on ice for 20 minutes. Cells were sonicated at medium intensity for 15 minutes (Bioruptor XL; Diagenode), and lysates were cleared by centrifugation at 16,000g for 20 minutes at 4ºC. 100 μL of the lysate were retained for inputs. 4 mg of the total protein extracts were incubated with 100 μL of pre-washed Strep-Tactin superflow resin (IBA GmbH, Gottingen, Germany) overnight at 4ºC using an overhead tumbler. The resin was washed 3 times with 500 μL wash Buffer (30 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). The proteins were eluted from the Strep-Tactin matrix in 50 μL of Elution buffer (5x Desthiobiotin Elution buffer (IBA GmbH) in TBS buffer (30 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.1% NP-40, 1X protease inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)) for 10 minutes on ice.

Western blotting analysis

For the immunoprecipitations, 40 μg of input protein and 12 μL of Strep-IP elution, with 6X SDS (0.2% Bromophenol blue and β-mercaptoethanol), were

separated by SDS-PAGE and transferred to 0.2 μm pore Nitrocellulose membranes (Amersham Protran; Sigma-Aldrich), except when probing for CHD7 and CHD8 that 0.45 μm pore Nitrocellulose membranes (Amersham Protran; Sigma-Aldrich) were used, and for RIF1 and Streptavidin, PVDF membranes (0.45 μm pore, Immobilon-P, Merck) were used. Membranes were blocked and antibodies prepared in 5% non-fat milk in PBS-T, with the exception of CHD7 and CHD8, where 5% BSA in PBS-T was used. The primary antibodies (Table 10) were detected with the appropriate secondary antibodies conjugated to Horseradish peroxidase (HRP) (Table 11) and visualized by ECL-Plus (GE Healthcare). In the case of BioID-WB, 40 μg of input protein and 10 μL of Strep-IP elution were used.

Table 10. Primary antibodies used in this study.

Antigen	Species	Source & reference	Dilution
TLK1	Rabbit	Cell Signaling #4125	1:1000 (WB)
TLK2	Rabbit	Bethyl Laboratories A301-257A	1:1000 (WB)
ASF1	Rabbit	Groth Laboratory (Groth et al., 2005)	1:2000 (WB)
LC8	Rabbit	Abcam ab51603, clone EP1660Y	1:1000 (WB)
RIF1	Rabbit	Bethyl Laboratories A300-568A	1:5000 (WB)
CHD7	Rabbit	Bethyl Laboratories A301-223A	1:2000 (WB)
CHD8	Rabbit	Bethyl Laboratories A301-224A	1:2000 (WB)
BRD4	Rabbit	Abcam ab128874	1:1000 (WB)
MTHFD2	Rabbit	Abcam ab151447	1:2000 (WB)
Rad50	Mouse	GeneTex GTX70228, clone 13B3	1:5000 (WB)
MRE11	Mouse	Santa Cruz sc-135992	1:5000 (WB)
NBS1	Mouse	GeneTex GTX70224, clone 1D7	1:5000 (WB)
Actin	Mouse	Sigma-Aldrich A4700	1:1000 (WB)
Lamin B1	Mouse	Santa Cruz sc-374015, clone B-10	1:1000 (WB)
GAPDH	Mouse	Santa Cruz sc-47724, clone 0411	1:1000 (WB)
FLAG	Mouse	Sigma-Aldrich F3165, clone M2	1:5000 (WB)
Strep-tag	Mouse	IBA GmbH 2-1509-001	1:1000 (WB)
M2 Flag	Mouse	Sigma-Aldrich F1804	1:500 (IF)
Lamin A	Rabbit	Santa Cruz sc-20680	1:500 (IF)
PyMT	Rat	Santa Cruz sc-53481	1:150 (IHC)
Iba-1	Rabbit	Wako #019-19741	1:1000 (IHC)
Ki67	Rabbit	Abcam ab15580	1:1000 (IHC)
CD45	Rat	Thermo Fisher Scientific 14-0451-82, clone 30-F11	1:100 (IHC)
Cleaved caspase 3	Rabbit	Cell Signaling #9661, clone Asp175	1:300 (IHC)
H2A.X-pS139	Rabbit	Cell Signaling #9718, clone 20E3	1:750 (IHC)
Rat IgG	Rat	R&D Systems, 6-001-F (isotype)	(IHC)
Rabbit IgG	Rabbit	Abcam, ab27478 (isotype)	(IHC)

WB, western blot; IF, immunofluorescence; IHC, immunohistochemistry; Nitro, nitrocellulose.

Table 11. Secondary antibodies used in this study.

Antigen	Source & reference	Dilution
Protein A/G anti-rabbit HRP	Thermo Fisher Scientific 32490	1:15000 (WB Nitro)
		1:30000 (WB PVDF)
goat anti-mouse IgG HRP	Thermo Fisher Scientific 31430	1:15000 (WB Nitro)
		1:30000 (WB PVDF)
Alexa Fluor 488 goat anti-mouse IgG	Thermo Fisher Scientific A28175	1:500 (IF)
Alexa Fluor 594 goat anti-rabbit IgG	Thermo Fisher Scientific A11012	1:500 (IF)
OmniMap anti-rat HRP	Roche 760-4457	1:500 (IHC)
Rabbit Anti-rat	Vector, AI-4001	1:500 (IHC)
OmniMap anti-rabbit HRP	Roche, 760-4311	1:500 (IHC)
BrightVision poly HRP anti-rabbit IgG	ImmunoLogic, DPVR-110HRP	1:500 (IHC)

WB, western blot; IF, immunofluorescence; IHC, immunohistochemistry; Nitro, nitrocellulose.

In vitro kinase assays from cell lysates pull-downs

In vitro kinase assays were performed after Strep-IP purification of pcDNA3.1 N-SF-TAP TLK1 and TLK2 coming from AD-293 cells. 200 μg of Strep-IP were incubated with 2 μCi ^{32}P-γ-ATP, 100 μM cold ATP, 1 μg of substrate in 12 μL of kinase buffer (50mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 2 mM DTT, 1X protein inhibitor cocktail (Roche) and 1X phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). The reaction was incubated at 30ºC for 30 minutes. After that, the reaction was stopped by adding 4 μL of Sample Buffer (6x SDS, (0.2% bromophenol blue and *β*-mercaptoethanol), and boiled for 5-10 minutes at 95ºC. Samples were analyzed on SDS-PAGE, stained with Coomassie Blue for 1 hour, washed 4 times with destaining buffer (10% acetic acid, 40% methanol, and 50% H_2O) and vacuum dried with an SGD2000 (Savant) for 2 hours at 60ºC. TLK1, TLK2, and substrate phosphorylation were measured using a Typhoon 8600 Variable Mode Imager (Molecular Dynamics) and band intensity quantified using ImageJ (Schindelin et al., 2012). Substrates used were 1 μg of Myelin basic protein (MBP) (Sigma-Aldrich) and 1 μg of ASF1a protein (kindly provided by Anja Groth).

When treating with inhibitors, the Strep-IP elute containing the purified kinase was pre-incubated in kinase buffer and cold-ATP with the inhibitors at the desired concentration for 30 minutes at room temperature, and then added the ^{32}P-γ-ATP and the substrate and proceeded with the regular reaction. The inhibitors tested in *in vitro* kinase assays are summarized in Table 12.

Table 12. Inhibitors tested in *in vitro* kinase assays.

Inhibitor	Concentration	Source	Reference
Staurosporine	10 μM	539745 PKi library II EMB Calbiochem	(Gao et al., 2013)
Thioridazine hydrochloride (THD)	16 μM	R750603 Sigma-Aldrich	(Ronald et al., 2013)
Fluphenazine dihydrochloride (FF)	20 μM	F4765 Sigma-Aldrich	
TLKi #1	10 μM	This study (N. Gray lab)	N. Gray collaboration
TLKi #2	10 μM	This study (N. Gray lab)	N. Gray collaboration
TLKi #3	10 μM	This study (N. Gray lab)	N. Gray collaboration

Alamar Blue viability assay

For viability assays, we first determined the plating density and incubation time of the Alamar Blue reagent (DAL1100, ThermoFisher Scientific). 24 hours after plating 700 U-2-OS cells/well in a 96 well plate, we treated the cells with the inhibitors at the concentrations indicated in Table 13. Viability was measured every 72 hours. Alamar Blue reagent was added diluted 1:10 to the cell culture media 5 hours prior to reading, and fluorescent data was collected using a FL600 microplate fluorescence reader (Bio-Tek). A Bliss test of independence was used to estimate synergistic activity between two drugs (Liu et al., 2018).

Table 13. Inhibitors used in viability assays.

Inhibitor	Concentration	Source
TLKi #1	500 nM	This study (N. Gray lab)
TLKi #3	20 nM	This study (N. Gray lab)
Chk1i	50 nM	PF-477736 – Selleckchem
ATRi	1 μM	ETP-46464 – Selleckchem
Olaparib	500 nM	AZD2281 – Selleckchem

Immunofluorescence (IF)

AD-293 cells were seeded and transiently transfected with the indicated plasmids. The next day, cells were trypsinized and seeded on poly-L-Lysine-coated coverslips. 48 hours post-transfection, cells were fixed with 4% formaldehyde (Santa Cruz Biotechnology) for 10 minutes and permeabilized in 0.2% Triton X-100 (Sigma-Aldrich) in 1x PBS for 10 minutes at room temperature. Coverslips were washed twice with PBS and blocked with PBS-BT (0.1% Triton X-100, 3% BSA (Sigma-Aldrich) in PBS) for 30 minutes at room temperature. The coverslips were incubated with the corresponding primary antibodies (Table 10) for 4 hours at 4ºC in a humid chamber. After three washes with PBS-BT, the coverslips were incubated with the secondary antibody (Table 11) for 1 hour at room temperature in a dark humid chamber. The coverslips were washed 3 times in PBS-BT and 4',6-diamidino-2-phenylindole (DAPI) was added diluted 1:3000 in the first wash. Fluorescent images were acquired with an Orca AG camera (Hamamatsu) mounted on a Leica DMI6000B microscope equipped with 1.4 numerical aperture 100X oil immersion objective. The phenotypic distribution was quantified in 10 different fields of view in each condition.

Proximity-dependent biotin identification mass spectrometry (BioID-MS)

AD-293 cells were seeded in 15 cm plates and transiently transfected the next day with 20 μg of BirA* plasmids using PEI and 150 mM NaCl. Medium was changed 6-8 hours post-transfection. 24 hours post-transfection, 50 μM of biotin (IBIAN Biotechnology; 2-1016-002) were added per plate. For mass spectrometry, 5x 15 cm plates were used per condition. 48 hours post-transfection, the cells were harvested with Trypsin-EDTA (Sigma-Aldrich) and the 5 plates per condition were pooled together. Cell pellets were washed twice in cold PBS and lysed in 5 mL of cold lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.1% SDS, 2 mM Mg_2Cl, 1% Triton X-100 (Sigma-Aldrich), 1 mM EDTA (Sigma-Aldrich), 1 mM EGTA (Sigma-Aldrich), 1:2000 benzonase 25 U/mL (Sigma-Aldrich), 1x protease inhibitor cocktail (Roche) and 1x phosphatase inhibitor cocktails 2&3 (Sigma-Aldrich)). 100 μL of the lysate were retained for Western blotting analysis. The remaining lysate was incubated with streptavidin-sepharose beads (GE Healthcare 2-1206-010) during 3 hours in an end-over-end rotator at 4ºC in order to isolate the biotinylated proteins. The beads were washed once in lysis buffer and three times in 50 nM ammonium bicarbonate pH 8.3 buffer. Samples were

snap-frozen and sent to the Mass Spectrometry & Proteomics Core Facility at IRB Barcelona for digestion and further analysis.

Tryptic digestion was performed directly on beads by incubation with 2 μg of trypsin in 50 mM NH_4HCO_3 at 37ºC overnight. The next morning, an additional 1 μg of trypsin was added and incubated for 2 hours at 37°C. The digestion was stopped by adding formic acid to 1% final concentration. Samples were cleaned through C18 tips (polyLC C18 tips) and peptides were eluted with 80% acetonitrile, 1% TFA. Samples were diluted to 20% acetonitrile, 0.25% TFA, loaded into strong cation exchange columns (SCX) and peptides were eluted in 5% NH_4OH, 30% methanol. Finally, samples were evaporated to dry, reconstituted in 50 μL and diluted 1:8 with 3% acetonitrile, 1% formic acid aqueous solution for nanoLC-MS/MS analysis.

The nano-LC-MS/MS was set up as follows. Digested peptides were diluted in 3% ACN/1% FA. Sample was loaded to a 300 μm × 5 mm PepMap100, 5 μm, 100 Å, C18 μ-precolumn (Thermo Scientific) at a flow rate of 15 μl/minute using a Thermo Scientific Dionex Ultimate 3000 chromatographic system (Thermo Scientific). Peptides were separated using a C18 analytical column Acclaim PEPMAP 100 75 μm x50 cm nanoviper C18 3 μm 100A (Thermo Scientific) with a 90 minute run, comprising three consecutive steps with linear gradients from 3 to 35% B in 60 minutes, from 35 to 50% B in 5 minutes, and from 50% to 85% B in 2 minutes, followed by isocratic elution at 85% B in 5 minutes and stabilization to initial conditions (A= 0.1% FA in water, B= 0.1% FA in CH_3CN). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos™ Tribrid (Thermo Scientific). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by CID and detected in the linear ion trap. The ion count target value was 400,000 and 10,000 for the survey scan and for the MS/MS scan respectively. Target ions already selected for MS/MS were dynamically excluded for 15 seconds. Spray voltage in the NanoMate source was set to 1.60 kV. RF Lens were tuned to 30%. Minimal signal required to trigger MS to MS/MS switch was set to 5,000. The spectrometer was working in positive polarity mode and single charge state precursors were rejected for fragmentation.

We performed a twin database search with two different softwares, Thermo Proteome Discoverer v2.3.0.480 (PD) and MaxQuant v1.6.6.0 (MQ). The search

engine nodes used were Sequest HT for PD and Andromeda for MQ. The databases used in the search was SwissProt Human (release 2019 01) including contaminants and TLK1 and TLK2 proteins. We run the search against targeted and decoy databases to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in M and acetylation in protein N-terminus as dynamic modifications. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.6 Da. Peptide were filtered at a false discovery rate (FDR) of 1 % based on the number of hits against the reversed sequence database.

For the quantitative analysis, contaminant identifications were removed and unique peptides (peptides that are not shared between different protein groups) were used for the quantitative analysis with SAINTexpress-spc v3.6.1 (Teo et al., 2014). SAINTexpress compares the prey control spectral counts with the prey test spectral counts for all available replicates. For each available bait and for each available replicate, we took as prey count the maximum count result between PD and MQ. Once obtained this combined dataset, we ran the SAINTexpress algorithm with TLK2 samples and a number of controls samples from previous experiments in the same cell type (n=45 total). High confidence interactors were defined as those with a SAINT score of 0.7 or greater. Output data from SAINTexpress is available upon request and published raw data for TLK2-WT, TLK2-D551G, and TLK2-S617L is available in the PRIDE repository, accession number PXD019450 (Pavinato et al., 2020).

Small-scale BioID validation by western blot (BioID-WB)

For the validation of the BioID results, we transfected AD-293 cells with 20 μg of BirA* as previously explained (3x 15 cm plates per condition). Biotin was added 24-hours post-transfection and samples harvested in the same way as for a BioID-MS experiment, with the only difference that after the ammonium bicarbonate washes we added 125 μL of 1X SDS to the beads and boiled them for 10 minutes at 96ºC. After this we centrifuged and kept the supernatant for western blot analysis.

Phospho-proteomics

A 15 cm plate of hTERT-RPE-1 cells was silenced with siGFP or siTLK1 + siTLK2 for 48 hours as previously explained. Cells were then lysed in 500 μL of 4% SDS, 0.1 M DTT in Tris-HCl pH 7.6. Samples were sonicated for 15 minutes

at medium intensity in 30" on - 30" off cycles. Lysates were then centrifuged at 16,000 g at 4ºC for 15 minutes. Supernatants were placed in a clean Eppendorf. 1 µL of benzonase was added and tubes were placed in an end-over-end rotor at 4ºC for 30 minutes.

Samples were quantified by Pierce 660 nm Protein Assay and Ionic Detergent Compatibility Reagent. 500 µg of each sample were digested in FASP with Trypsin overnight at 37ºC following FASP protocol (Wisniewski et al., 2009). Digested samples were acidified to a final concentration of 1% formic acid. Then, samples were cleaned up through polyLC C18 tips, peptides were eluted with 80% acetonitrile, 1% TFA and evaporated to dryness. Samples were reconstituted in 150 µL of Binding/Equilibration buffer from High-Select™ TiO2 Phosphopeptide Enrichment Kit (Thermo Fisher Scientific #A32993) and enriched in phosphopeptides using this kit. Phosphopeptides were eluted with Phosphopeptides Elution Buffer and acidified with formic acid. Samples were evaporated to dryness and then reconstituted in 25 µL of 3% acetonitrile, 1% formic acid aqueous solution MS analysis. Samples were loaded to a 300 µm x 5 mm PepMap100, 5 µm, 100 Å, C18 µ-precolumn (Thermo Fisher Scientific) at a flow rate of 15 µL/minute using a Dionex Ultimate 3000 chromatographic system (Thermo Fisher Scientific). Peptides were separated using a C18 analytical column (NanoEase MZ HSS T3 column (75 µm x 250 mm, 1.8 µm, 100A), Waters) with a 180 minute run, comprising three consecutive steps with linear gradients from 3 to 35% B in 150 minutes, from 35 to 50% B in 5 minutes, and from 50 % to 85 % B in 1 minute, followed by isocratic elution at 85 % B in 3 minutes and stabilization to initial conditions (A= 0.1% formic acid in water, B= 0.1% formic acid in CH_3CN). The column outlet was directly connected to an Advion TriVersa NanoMate (Advion) fitted on an Orbitrap Fusion Lumos™ Tribrid (Thermo Fisher Scientific). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the Orbitrap with the resolution (defined at 200 m/z) set to 120,000. The lock mass was user-defined at 445.12 m/z in each Orbitrap scan. The top speed (most intense) ions per scan were fragmented by HCD and detected in the Orbitrap. The ion count target value was 400,000 and 10,000 for the survey scan and for the MS/MS scan respectively. Target ions already selected for MS/MS were dynamically excluded for 15 seconds. Spray voltage in the NanoMate source was set to 1.60 kV. RF Lens were tuned to 30%. Minimal signal required to trigger MS to MS/MS switch was set to 5,000. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation.

For non-enriched samples, a database search was performed with Proteome Discoverer software v2.3.0.480 (Thermo Fisher Scientific) using Sequest HT search engine and SwissProt Human released 2019_01 with contaminants database. Search was run against targeted and decoy database to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in M and acetylation in protein N-terminus as dynamic modifications, and carbamidomethylation in C as static modification. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.6 Da. Peptides with a q-value lower than 0.1 and a FDR < 1% were considered as positive identifications with a high confidence level. For enriched samples, a database search was performed with Proteome Discoverer software v2.3.0.480 (Thermo Fisher Scientific) using Sequest HT search engine, MS Amanda 2.0 search engine and SwissProt Human released 2019_01 with contaminants database. Search was run against targeted and decoy database to determine the false discovery rate (FDR). Search parameters included trypsin enzyme specificity, allowing for two missed cleavage sites, oxidation in M, acetylation in protein N-terminus and phosphor in STY as dynamic modifications, and carbamidomethylation in C as static modification. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.02 Da. Peptides with an FDR < 1% were considered as positive identifications with a high confidence level.

Web Resources

Table 14. Web resources used in this thesis.

Resource	Domain
Autism Sequencing Consortium exome analysis browser	https://asc.broadinstitute.org/
BioGRID Database	https://thebiogrid.org/
ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/
Constrained Coding Regions Browser	https://s3.us-east-2.amazonaws.com/ccrs/ccr.html
DECIPHER	https://decipher.sanger.ac.uk/
gnomAD Browser, v2.1.1	https://gnomad.broadinstitute.org/
Modeller, v9.25	https://salilab.org/modeller/
OMIM	https://omim.org/
RepeatMasker Open-4.0.	http://www.repeatmasker.org
SFARI gene	https://gene.sfari.org/
STRING Database	https://string-db.org/
Structuropedia web interface	http://mod.farooq.ac/

Varsome	https://varsome.com/
Worldwide Protein Data Bank	https://www.wwpdb.org/

Results

Chapter 1

Tousled like kinases as a target in cancer therapy

Marina Villamor-Payà[1], Camille Stephan-Otto Attolini[1], Adrià Caballé Mestres[1], Neus Prats[1], Lluís Palenzuela[1], Lars Koenig[2], Mohamed-Reda Benmebarek[2], Travis H. Stracker[1, 3]

[1]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain

[2]Division of Clinical Pharmacology, University Hospital, LMU Munich, 80337 Munich, Germany

[3]Radiation Oncology Branch, National Cancer Institute, Bethesda, MD 20892, USA

Statement of contribution:

M Villamor-Payà designed, performed and analyzed all experiments.

C Stephan-Otto Attolini analyzed TCGA datasets in Figure R1 with feedback from

A Caballé Mestres performed bioinformatic analysis of Figures R3A and R4C.

N Prats performed histological analysis of lung metastasis shown in Figure R4.

L Palenzuela helped maintain mouse cohorts and with mouse genotyping.

L Koenig designed and provided antibodies for the FACS panels in Figures R6 and R7.

M-R Benmebarek helped analyze FACS panels in Figures R6 and R7.

TH Stracker supervised the project.

TLK expression anti-correlates with immune signaling in many human cancers

Solid tumors do not solely consist of tumor cells, but also of fibroblasts, blood vessels and immune cells, as well as an extracellular matrix (ECM) containing a soup of extracellular molecules, such as cytokines and growth factors. This mix of cells and molecules is known as the tumor microenvironment (TME) (Binnewies et al., 2018). It has become clear that while driver mutations are important for cancer progression and response to therapy, the TME also plays a crucial role in cancer development and avoidance of immunosurveillance (Duan et al., 2020). Depending on the composition of the TME, tumors can be placed into two distinct categories: *cold tumors*, which have low or undetectable lymphocyte infiltration, and *hot tumors*, that have high T cell infiltration, accompanied by high levels of pro-inflammatory cytokines (Gajewski, 2015). In general, hot tumors have a better prognosis, as they respond to immune checkpoint blockade (ICB) therapies such as anti-programmed death ligand (PD-L)1/ Programmed cell death protein (PD)-1 treatment (Gajewski, 2015, Duan et al., 2020, Zemek et al., 2020). Efforts in developing novel therapies that enable ICB effectiveness in patients with cold tumors are being made (Petitprez et al., 2020, Zemek et al., 2020).

We found that TLK depletion *in vitro* leads to the production of extra-telomeric DNA that is sensed by the cGAS-STING pathway that triggers a strong innate immune response (Segura-Bayona et al., 2020). cGAS-STING is known to be the main cytosolic DNA sensor, and can be activated by numerous sources of immunostimulation stemming from genome unstable tumors, including replication fork byproducts and micronuclei (Mackenzie et al., 2017, Coquel et al., 2018). cGAS-STING contributes to the effective production of Type I IFN and priming of cytotoxic T cells against the tumor-associated antigens (Duan et al., 2020). Tumors often silence STING to minimize immune surveillance (Wu et al., 2018) and thus STING agonists are currently being developed and trialed to boost cancer immunotherapy (Le Naour et al., 2020).

As high levels of TLK expression correlate with poor prognosis in several human cancer types (Kim et al., 2016, Lee et al., 2018), we wondered if this could be due not only to the protection provided from replication stress (Lee et al., 2018) but

also to changes in STING levels or the immune infiltration status of those tumors. To address this, we analyzed publicly available datasets from The Cancer Genome Atlas (TCGA) for correlations between *TLK1*, *TLK2*, STING (encoded by the *TMEM173* gene) levels, genome instability and immunity across cohorts of different tumor types. We observed that high levels of *TLK1* and *TLK2* correlated with low *STING* levels across many cohorts (Figures R1A and R1B). This indicates that tumors with unstable genomes could be selecting for STING silencing and high levels of TLK to suppress toxic levels of replication stress and avoid immune detection. Next, we examined chromosomal instability (CIN) levels using a previously published signature (Carter et al., 2006) and aneuploidy score, estimated from comparative genomic hybridization (CGH) SNP arrays (Taylor et al., 2018). *TLK2* expression, and to a lesser extent *TLK1* expression, correlated with CIN and aneuploidy levels in most tumor types (Figure R1A).

We next addressed whether tumors with high levels of TLK expression were differentially infiltrated by the immune system. We addressed the leukocyte and stromal fractions based on methylation (Thorsson et al., 2018) and expression (Aran et al., 2017) signatures, as was previously carried out by Taylor and colleagues (Taylor et al., 2018). We found several tumor types where high TLK2 expression strongly correlated with reduced immune infiltration (Figure R1A), such as low-grade glioma (LGG), liver hepatocellular carcinoma (LIHC) (Figure R1B), adrenocortical carcinoma (ACC), glioblastoma multiforme (GBM), and sarcoma (SARC). Interestingly, these cancer subtypes have been previously described as cold or immuno-depleted tumors (Thorsson et al., 2018).

With CIN, STING levels and immune infiltration, the correlation was stronger in the case of TLK2, which is also more frequently amplified in cancer (Lee et al., 2018). These results showed that tumors with high TLK2 expression appeared to exhibit decreased cytosolic DNA sensing by STING in order to avoid being recognized by the immune system. By treating patients with TLK high tumors with TLK inhibitors, we might be able to increase immune cell infiltration, thereby 'heating up' their cold tumors and improving their prognosis and overall survival.

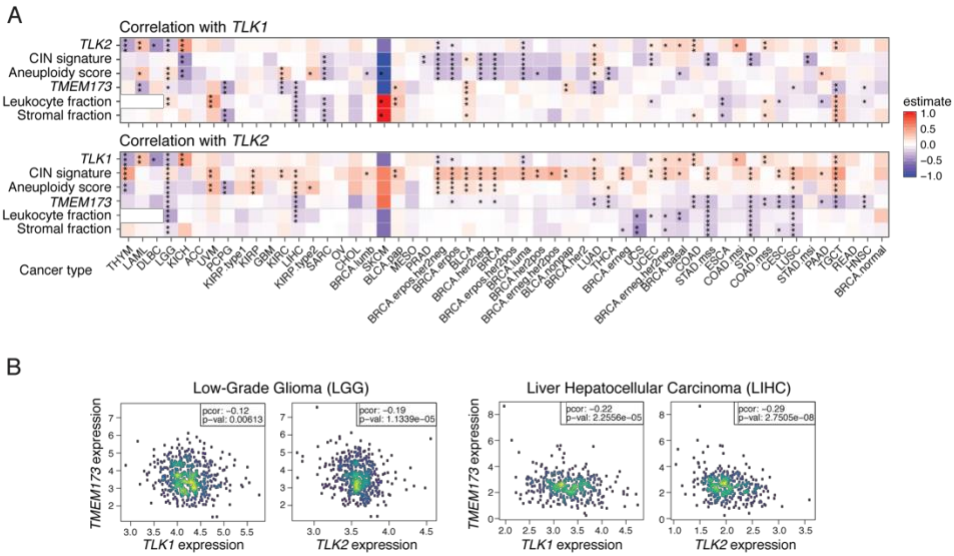


Figure R1. TLK expression correlates with suppressed immune signaling in human cancers.

(**A**) Heatmaps with correlations of expression levels of *TLK1* and *TLK2* with CIN signature, aneuploidy score, STING (*TMEM173*) levels, and signatures of tumor-immune populations across different TCGA datasets. Stars indicate adjusted p-values for multiple comparisons using Benjamini-Hochberg (***p < 0.001, **p< 0.01, *p < 0.05). Information on cancer types and number of patients per group are available in Supplementary Table 1. (**B**) Correlation between *TLK1* and *TLK2* and STING (*TMEM173*) in the TCGA datasets of low-grade glioma (LGG) and liver hepatocellular carcinoma (LIHC).

Establishing a MMTV-PyMT mouse model to assess the role of TLK2 in breast cancer progression

Prior attempts were made in the lab to assess the role of TLK1 and TLK2 in cancer progression *in vivo*. First, a xenograft system was set up in which parental MDA-MB-231 WT and MDA-MB-231 TLK1 knockout single clones were injected into athymic Balb/c mice. In those initial experiments, depleting TLK1 had an unexpectedly strong effect in tumor growth. However, when injecting pools of MDA-MB-231 WT and TLK1 knockout clones, it became evident that heterogeneity among single cell clones was high and a clear phenotype related to TLK1 loss was not evident. Additional efforts to generate a stable TLK2

knockout MDA-MB-231 line were unsuccessful (Segura-Bayona, 2018). Therefore, we decided to generate a genetic mouse model to address the effect of depleting TLK2 in breast cancer tumors *in vivo*.

The MMTV-PyMT model is a well-established transgenic mouse system in which mice carry an allele of the Polyomavirus Middle T antigen (PyMT) under the transcriptional control of the mouse mammary tumor virus (MMTV) promoter in the mammary gland (Guy et al., 1992). This results in the development of multifocal mammary tumors that metastasize almost exclusively to the lungs. The MMTV-PyMT model recapitulates the progression of human breast cancers with poor prognosis. Initially, the tumors exhibit high expression of estrogen and progesterone receptors (ERs and PRs), which are lost over time, in conjunction with ErbB2/Neu and cyclin D1 overexpression (Lin et al., 2003). This mouse model is characterized by a short latency and a high penetrance. We decided to use C57BL/6 immunocompetent mice to address the role of the immune system in tumor progression (Oh et al., 2014), knowing that the latency is increased in comparison to other genetic backgrounds (Davie et al., 2007). In our hands, measurable mammary tumors were detected at 14-20 weeks of age.

To generate MMTV-PyMT-expressing *Tlk2*$^{F/F}$ UbiquitinC-Cre^{ERT2} (UbCre^{ERT2}) mice, we crossed our previously established TLK2$^{F/F}$ mice (Segura-Bayona et al., 2017) with MMTV-PyMT (Guy et al., 1992) and UbCre^{ERT2} mice (Gruber et al., 2007) (Figure R2A). Experiments were initiated when the primary tumor reached a volume of 150-200 mm^3. Mice were then treated for 5 consecutive days via intraperitoneal injection of 4-Hydroxytamoxifen (4-OHT) to induce UbCre^{ERT2} activation and TLK2 deletion (Figures R2B and R2C). After this, the tumors were measured, and the mice weighed 2-3 times a week and monitored for signs of discomfort (Figure R2B). Prior to the main experiments, we tested the efficiency of the system to knockout TLK2. We genotyped the mice when they were 3-weeks old and at the experimental endpoint. An example of the PCR-based genotyping system used is shown, where in a TLK2$^{F/F}$ UbCre^{ERT2} mouse we observe the 450 bp *Tlk2*$^{-}$ band (right) at the endpoint, while in a *Tlk2*$^{F/F}$ mouse we can only see the 1300 bp *Tlk2*^{F} band (left) (Figure R2D). To test if TLK2 protein was also depleted, we performed western blotting on tumor tissue dissections at the endpoint. This confirmed that TLK2 protein cannot be detected in the *Tlk2*$^{-/-}$ mouse and that it is not compensated for by an increase in TLK1 expression (Figure R2E). Thus, the MMTV-PyMT system represents a

genetically modified mouse model in which TLK2 can be efficiently knocked out at different stages of mammary tumorigenesis to monitor tumor progression and metastasis *in vivo*.

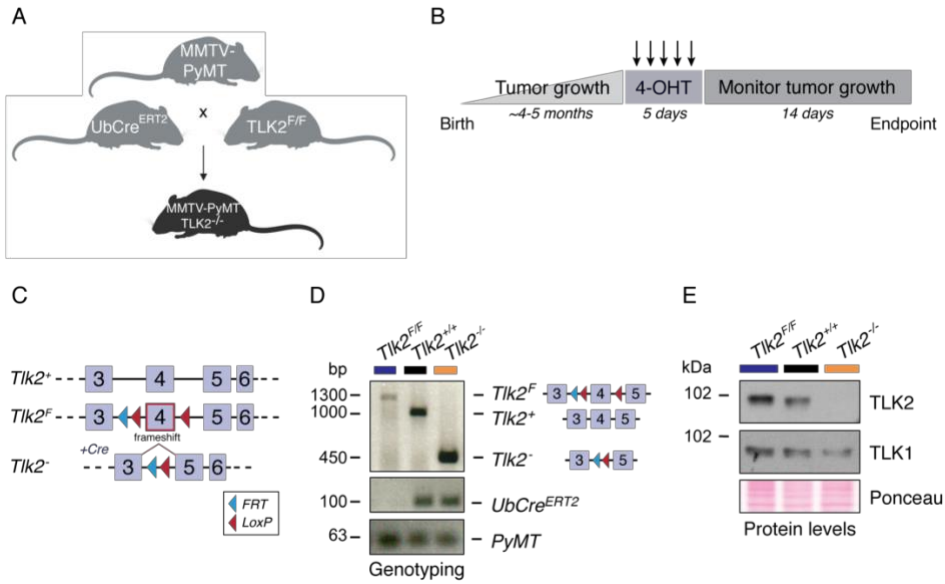


Figure R2. Establishment of a MMTV-PyMT mouse model to assess the role of TLK2 in breast cancer progression.

(**A**) In order to generate MMTV-PyMT-bearing *Tlk2*$^{F/F}$ UbCre^{ERT2} mice we crossed MMTV-PyMT (Guy et al., 1992), *Tlk2*$^{F/F}$ mice (Segura-Bayona et al., 2017) and UbCre^{ERT2} mice (Gruber et al., 2007). (**B**) Experimental layout. Mice were monitored for the appearance of spontaneous mammary tumors and treated with 5 intraperitoneal 4-OHT injections when the primary tumor reached a volume of 150-200 mm^3. Tumor size was then monitored for 14 days. (**C**) Schematic of the *Tlk2* wild-type (*Tlk2*$^{+}$), flox (*Tlk2*^{F}) and knockout (*Tlk2*$^{-}$) alleles. In the *Tlk2*^{F} allele, exon 4 is surrounded by two LoxP sites. In the presence of Cre recombinase, exon 4 is effectively deleted, resulting in a frameshift and premature termination of the mRNA and its degradation by the nonsense mediated mRNA decay pathway (NMD), generating the *Tlk2*$^{-}$ allele. **(D)** Example of the PCR-based genotyping of MMTV-PyMT^{+} *Tlk2*$^{+/+}$, *Tlk2*$^{F/F}$ and *Tlk2*$^{-/-}$ mice. DNA was extracted from tip-tails at endpoint. **(E)** Western blotting of TLK1 and TLK2 protein levels of the indicated genotypes from tumor tissue dissections at endpoint. Ponceau staining shows equal loading.

TLK2 loss stalls primary tumor growth

In order to properly assess the role of TLK2 loss in cancer progression using a UbCre^{ERT2} inducible system, we needed to control for both the effects of the Cre and of the 4-OHT treatment. The pro-drug tamoxifen is a widely used endocrine treatment for early ER^{+} breast cancers (Tremont et al., 2017). At the time of treatment with 4-OHT, most ER expression would have already been lost (Lin et al., 2003 and unpublished data by us and Nebreda's lab). Apart from this, the Mutant Estrogen Receptor (ERT2) that is fused to the UbCre recombinase binds 4-OHT with a very high affinity, allowing for controlled genetic recombination (Felker et al., 2016).

We measured and followed all of the tumors of each mouse but focused on the primary tumor, that is the first one that reaches a size of 150-200mm^3 before starting the 4-OHT treatment. As in every genetic mouse model, there was some variability between the mice of the same genotype (Figure R3A, thin lines). In general, we could see how most control tumors kept growing at a steady rate after 4-OHT treatment (*Tlk2*$^{F/F}$ and *Tlk2*$^{+/+}$ UbCre^{ERT2}), while the growth of the knockout tumors stalled (*Tlk2*$^{-/-}$ UbCre^{ERT2}) (Figure R3A, thick lines). Moreover, we could observe a relative reduction in tumor size in the *Tlk2*$^{F/F}$ mice when compared to *Tlk2*$^{+/+}$ UbCre^{ERT2} mice (Figure R3A). However, this latter observation was not statistically significant however.

We next addressed the proliferation status of these tumors by Ki67 immunohistochemical staining (Figures R3B and R3C). We previously observed that when we derived cell lines by spontaneous immortalization from MMTV-PyMT *Tlk2*$^{F/F}$ UbCre^{ERT2} tumors and induced TLK2 depletion *in vitro*, cells showed a decrease in colony formation capacity and reduced BrdU incorporation (Segura-Bayona, 2018). Although we observed a slight reduction in the proliferative capacity of *Tlk2*$^{-/-}$ UbCre^{ERT2} tumors (Figure R3C), it does not appear to be the only factor resulting in the observed reduction in tumor size.

Knocking out TLK2 in already formed MMTV-PyMT tumors stalled tumor growth but was not sufficient to completely clear the mice from cancer. These results suggest that treating patients with TLK inhibitors, especially those with TLK2 amplification, could slow down the progression of the disease. While

potentially curative, it may be more efficacious to target TLK2 in combination with other existing treatments. It is also worth noting that the MMTV-PyMT mouse is engineered to express the PyMT oncogene in all of the cells that activate the MMTV promoter. This would not reflect the *de novo* acquisition of an activated oncogene in relatively few cells of the mammary tissue, a more common scenario in human breast cancer. There are murine models such as the Replication-competent avian sarcoma-leukosis virus (RCAS)-MMTV system that allows for the activation of the oncogene only in a few cells of the mammary epithelium thus recapitulating better replication stress during early tumor initiation (Du et al., 2006, Gupta et al., 2013).

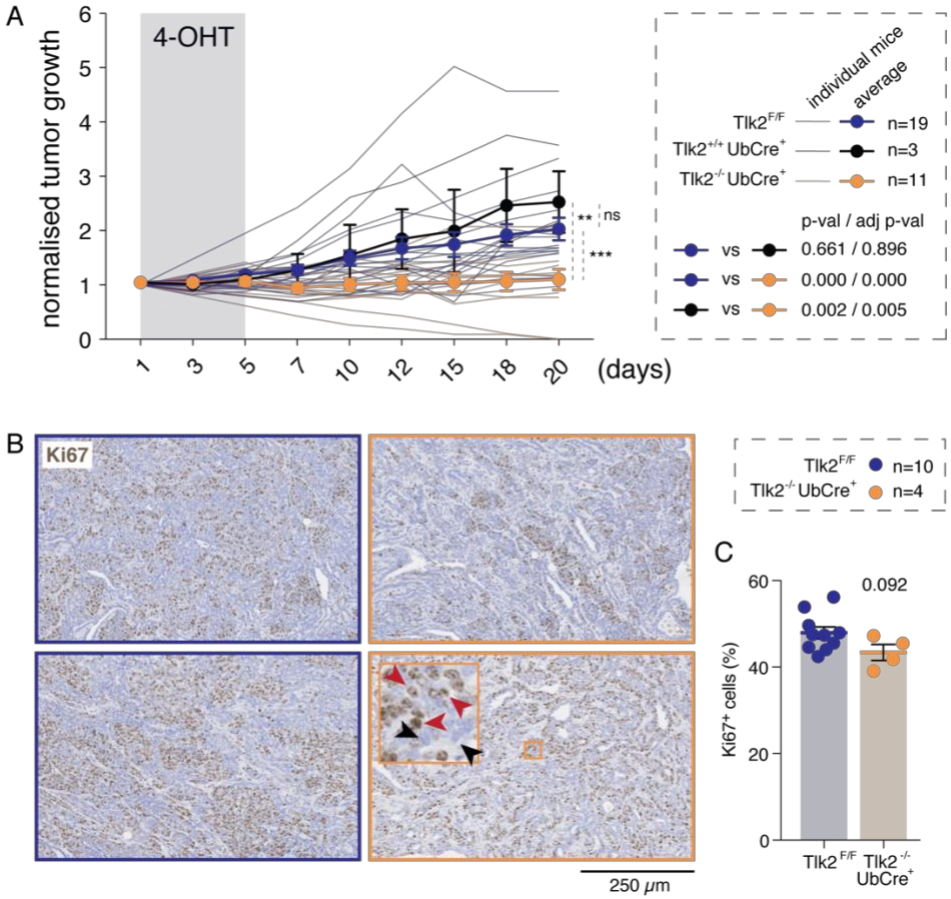


Figure R3. TLK2 loss stalls primary tumor growth.

(**A**) Normalized tumor growth curve showing all individual mice (thin lines) and average (thick lines). Tumor growth rates were compared between genotypes using a linear mixed effects model, R package (nlme) (Box, 1994, Pinheiro, 2000) with an autoregressive (AR-

1) block diagonal (indicating independence between subjects) covariance structure (Zavrakidis et al., 2020). The logarithm of the tumor volumes was considered as response variable, the time and the interaction between time and genotype were set as fixed effects, and the animal subject was set as random effect. Adjustment for multiple testing (single-step correction method) was performed using the R package multcomp (Hothorn et al., 2008). (***P<0.001, **P<0.01). For clarity of representation, data are shown in the original scale. MMTV-PyMT *Tlk2*$^{F/F}$ n=19, MMTV-PyMT *Tlk2*$^{+/+}$ UbCre^{ERT2} n=3, and MMTV-PyMT *Tlk2*$^{-/-}$ UbCre^{ERT2} n=11. (**B**) Representative images of Ki67 immunohistochemical staining performed in primary MMTV-PyMT tumors. Ki67 positive cells (red arrowheads) and Ki67 negative cells (black arrowheads) are indicated. (**C**) Quantification of (**B**). The percentage of Ki67^{+} cells was calculated using QuPath software (Bankhead et al., 2017). MMTV-PyMT *Tlk2*$^{F/F}$ n=11 and MMTV-PyMT *Tlk2*$^{-/-}$ UbCre^{ERT2} n=4. Error bars represent SEM and statistical significance was determined using an unpaired t test with Welch's correction.

TLK2 loss reduces lung metastasis incidence

MMTV-PyMT tumors metastasize to the lungs with a high incidence (Lin et al., 2003). To see if knocking out TLK2 had an effect on metastasis incidence or size, we collected the lungs of the tumor-bearing mice at the endpoint and performed a hematoxylin and eosin (H&E) staining and screened for metastasis. Macro metastasis can be seen as clusters of cells (dashed blue lines, Figure R4B), that are highly proliferative and have a dark and heterochromatic nucleus (black arrowheads Figure R4B). 45% of the control mice (*Tlk2*$^{F/F}$ and *Tlk2*$^{+/+}$ UbCre^{ERT2} mice, WT) developed macro metastasis of various sizes and in many cases accompanied by immune infiltrates (Figure R4C, red arrowheads Figure R4B). In the case of *Tlk2*$^{-/-}$ UbCre^{ERT2} (KO), we did not observe a single macro metastasis. In lungs from all genotypes without macro metastasis we observed some small clusters of cells that seemed compatible with micro metastasis, located mostly close to blood vessels (pink dashed lines, Figure R4D). In order to discriminate if these cells were tumor cells (PyMT^{+}) or immune cells (CD45^{+}), we set up a combined PyMT + CD45 staining, which revealed that these clusters are immune infiltrates (pink dashed lines, Figure R4E).

These results are in concordance with the tumor progression data (Figure R3A) and show that primary tumors that are more aggressive also metastasize more.

These results highlight that knocking out TLK2 does not only have an effect on primary tumor progression, but also on the incidence of lung metastasis, that is strikingly abolished (Figures R4A and R4C). These findings are very encouraging, as metastasis is the main cause of death for patients with breast cancer (Bertucci et al., 2019).

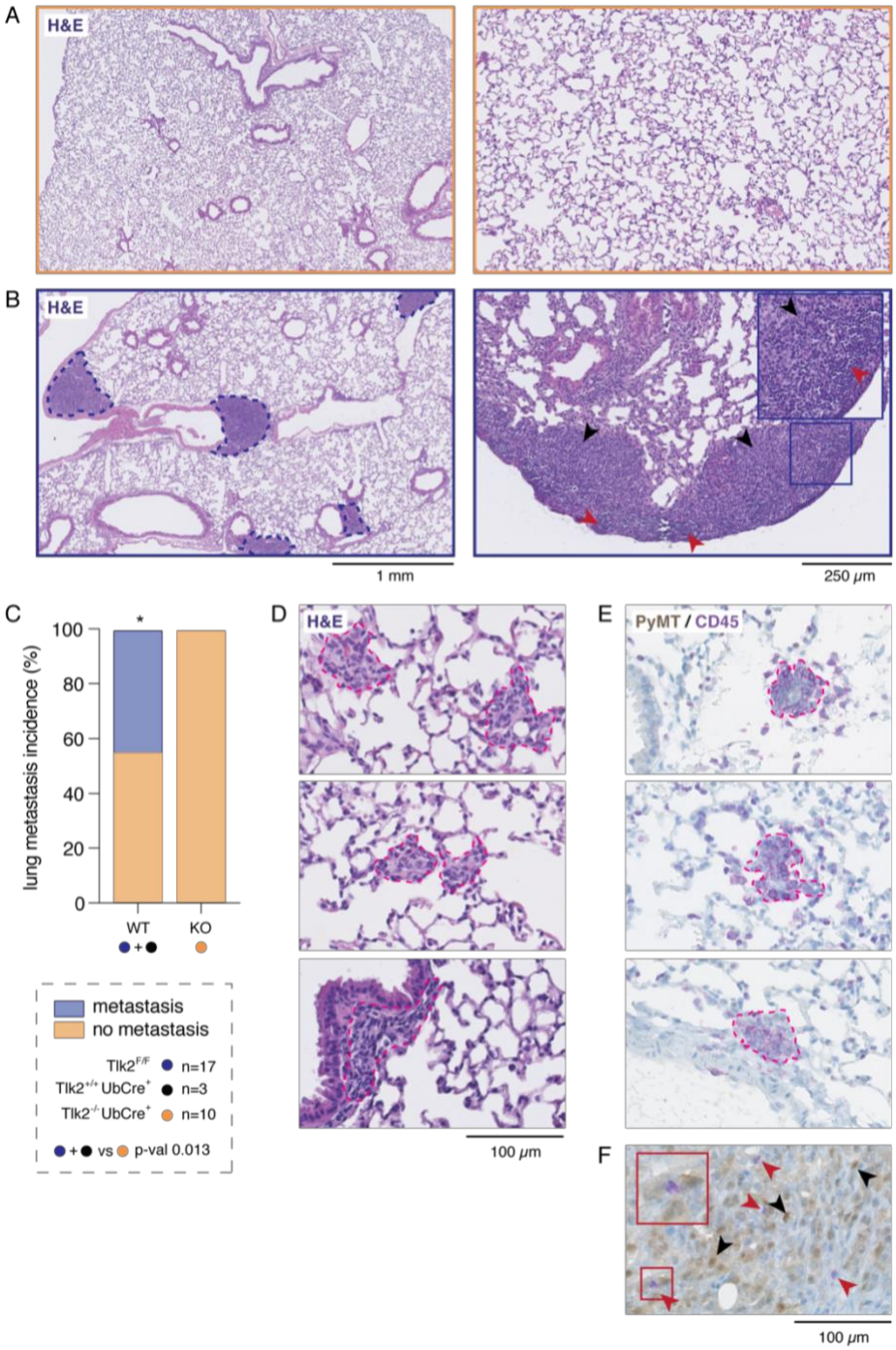


Figure R4. TLK2 loss reduces lung metastasis incidence.

(A) Representative images of Hematoxylin an eosin (H&E) immunohistochemical staining of healthy lung tissue from a MMTV-PyMT *Tlk2*$^{-/-}$ UbCre^{ERT2} mouse. **(B)** Representative images of H&E immunohistochemical staining of lung tissue from a MMTV-PyMT TLK2$^{F/F}$ mouse featuring metastasis. In the left panel, macro metastasis are indicated in dashed blue lines. In the right panel, an example of a metastasis accompanied by immune infiltrate is shown. Metastatic cells (black arrowheads) have a dark and heterochromatic nucleus while immune cells (red arrowheads) are smaller and have a very dark and compact nucleus. **(C)** Bar graph showing the percentage of mice that present with lung metastasis. Serial cuts of the lungs were performed, and 5 cuts/lung were blindly screened for metastasis. As there were no differences in primary tumor growth between *Tlk2*$^{F/F}$ and *Tlk2*$^{+/+}$ UbCre^{ERT2} mice, we decided to pool them together (WT) for metastasis analysis to gain statistical power. A Fisher's exact test of independence (Fisher, 1922) was performed (*P<0.05). MMTV-PyMT *Tlk2*$^{F/F}$ n=17, MMTV-PyMT *Tlk2*$^{+/+}$ UbCre^{ERT2} n=3, and MMTV-PyMT *Tlk2*$^{F/F}$ UbCre^{ERT2} n=10. **(D)** Representative images of H&E immunohistochemical staining of lung tissues showing potential lung micro metastasis (pink dashed lines). **(E)** Representative images of PyMT + CD45 immunohistochemical staining of lung tissues showing potential lung micro metastasis (pink dashed lines). PyMT (brown) stains for tumor cells while CD45 (purple) stains for immune cells. **(F)** Representative image of PyMT + CD45 immunohistochemical staining of a primary breast tumor. In brown are PyMT^{+} cells (black arrowheads) and in purple are CD45^{+} cells (red arrowheads).

Increased immune infiltration in TLK2 knockout tumors at early time points

MMTV-PyMT is an immune-permissive tumor model that has previously been used to study tumor-stroma interactions (Davie et al., 2007, Fluck and Schaffhausen, 2009). In order to try and understand what the driver of tumor reduction was in our setting, we analyzed the main cellular components of the TME (tumor cells, immune cells, cancer-associated fibroblasts, and endothelial cells) by flow cytometry (Figure R5). We did not observe a significant change in the percentage of the different cell types, regardless of TLK2 status.

We then performed microarrays of EpCam^{+} sorted cells from the tumors at the endpoint to determine if there was differential gene expression that could shed light on the tumor growth differences, but we again did not observe any significant differences. This could be due to the time-point or due to the fact that non-tumor cell autonomous factors influence tumor regression after TLK2 loss.

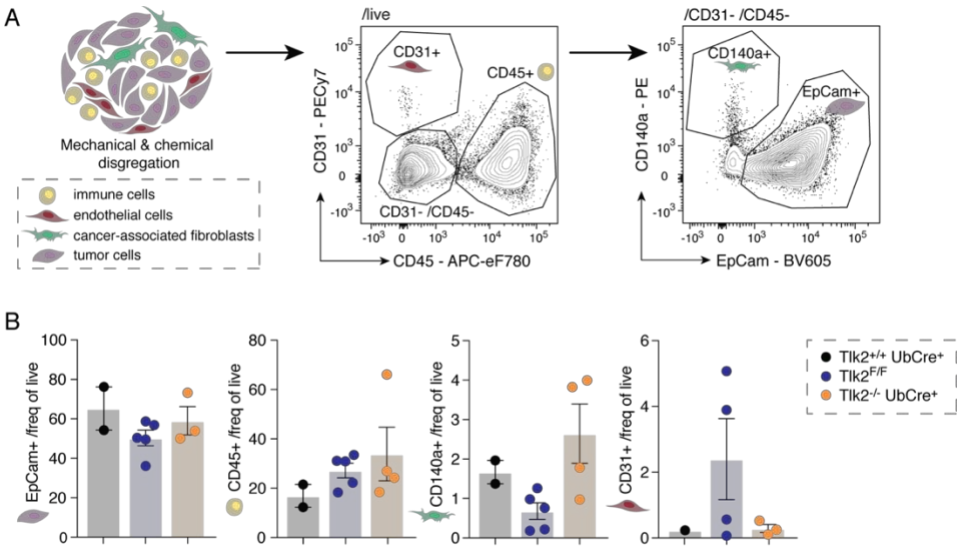


Figure R5. TLK2 wild-type and knockout tumors have a similar TME composition.

(**A**) Gating strategy used to study the TME populations within the tumor: tumor cells (EpCam$^+$), immune cells (CD45$^+$), cancer-associated fibroblasts (CD140$^+$), and endothelial cells (CD31$^+$). (**B**) Bar graphs showing the stated populations out of live cells. Error bars represent SEM. Statistical significance was determined using an unpaired t test with Welch's correction.

As we did not see any significant differences in major cell populations (Figure R5B), we set-up an exploratory flow cytometry analysis to better understand the different myeloid (Figure R6) and lymphoid (Figure R7) immune populations that infiltrate these tumors during the course of their development. We first analyzed tumors 15 days post 4-OHT treatment (as in Figures R2 and R3). At this timepoint we observed a tendency for reduced numbers in all circulating leucocytes: neutrophils (Figure R6F), monocytes (Figure R6I) and lymphocytes (Figure R7B). We also ran a small pilot experiment in which we sacrificed the mice only 3 days post 4-OHT treatment (Figure R6A). *Tlk2* knockout tumors were observably smaller at this early timepoint (Figures R6A and R6B). Readouts at 3- and 15-days post treatment allowed us to assess the early immune response and how it comparably changes during the course of tumor regression. We plotted the results from both time-points side by side to simplify the interpretation of the data. In most cases, this correlated well with an increase of

particular populations within the tumor at the earlier time-point, and a small decrease at the later time-point (Figures R6G, R6J and R7C).

We observed an early increase in monocytes and immature tumor-associated macrophages (TAMs) (Figures R6G and R6J) that matured and led to a late increase in TAMs (Figure R6K). Both monocytes and TAMs showed a slight increase in PD-L1 expression at 15 days post 4-OHT (Figures R6H and R6L). In the case of neutrophils, the small decrease observed at day 15 (Figure R6D) was accompanied with increased PD-L1 expression (Figure R6E).

In the lymphoid lineage, at the early time-point we observed an increase in the frequency of NK cells (Figure R7C). Although this increase was more modest in the $CD3^+$ T cell population (Figure R7D), the cells appeared more activated, with the emergence of a distinct $CD3^{high}$ subpopulation (Figures R7A and R7G). Expression of the early activation markers CD69 (Figure R7F) and PD-1 were also increased (Figure R7E).

Despite observable shifts in distinct immune populations, this dataset still lacks statistical power due to the amount of heterogeneity and low sample number. Interestingly, it appears to be congruent with the above-described TCGA dataset highlighting the anti-correlation between TLK expression and leucocyte and stromal fractions (Figure R1). Understanding the precise impact of knocking out TLK2 on tumor cells and its effects on immune cell function remains to be determined.

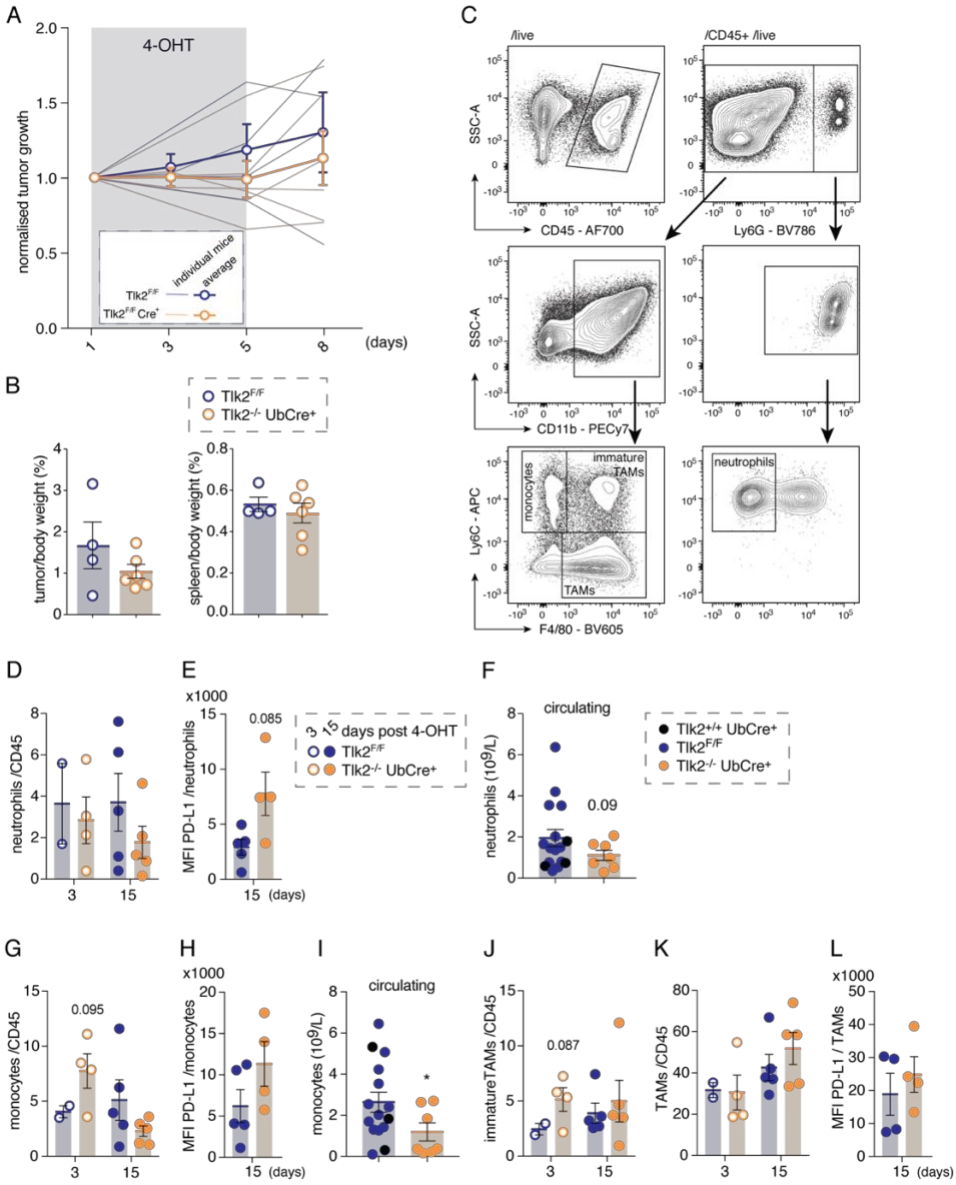


Figure R6. TLK2 knockout tumors present with an early increased myeloid infiltration.

(**A**) Normalized tumor growth curve showing all individual mice (thin lines) and average (thick lines) of mice sacrificed 3 days after the 4-OHT treatment finished. MMTV-PyMT *Tlk2*$^{F/F}$ n=4 and MMTV-PyMT *Tlk2*$^{-/-}$ UbCre^{ERT2} n=6. (**B**) Tumor and spleen weight normalized to body weight of the mice shown in (**A**). (**C**) Gating strategy used to study the myeloid populations within the tumor: neutrophils (CD45^{+}/Ly6G^{+}/CD11b^{+}/Ly6C^{-}/F4/80^{-}), monocytes (CD45^{+}/Ly6G^{-}/CD11b^{+}/Ly6C^{+}/F4/80^{-}), immature TAMs

(CD45$^+$/Ly6G$^-$/CD11b$^+$/Ly6C$^+$/F4/80$^+$), TAMs (CD45$^+$/Ly6G$^-$/CD11b$^+$/Ly6C$^-$/F4/80$^+$). (**D**, **G**, **J**, and **K**) percentage of the stated populations /CD45$^+$ cells. Neutrophils (**D**), monocytes (**G**), immature TAMs (**J**), and TAMs (**K**) at days 3 and 15 post 4-OHT treatment. (**E**, **H**, and **L**) Mean Fluorescence Intensity (MFI) of PD-L1 of the indicated populations at day 15 post 4-OHT treatment. Neutrophils (**E**), monocytes (**H**), and TAMs (**L**). (**F** and **I**) Circulating cells / liter (L) assessed by a complete blood count (CBC) performed at day 15 post 4-OHT treatment. Neutrophils (**F**) and monocytes (**I**). Error bars represent SEM. Statistical significance was determined using an unpaired t test with Welch's correction (*$P<0.05$) and comparing the genotypes at the same day post 4-OHT treatment. TAM; tumor-associated macrophage.

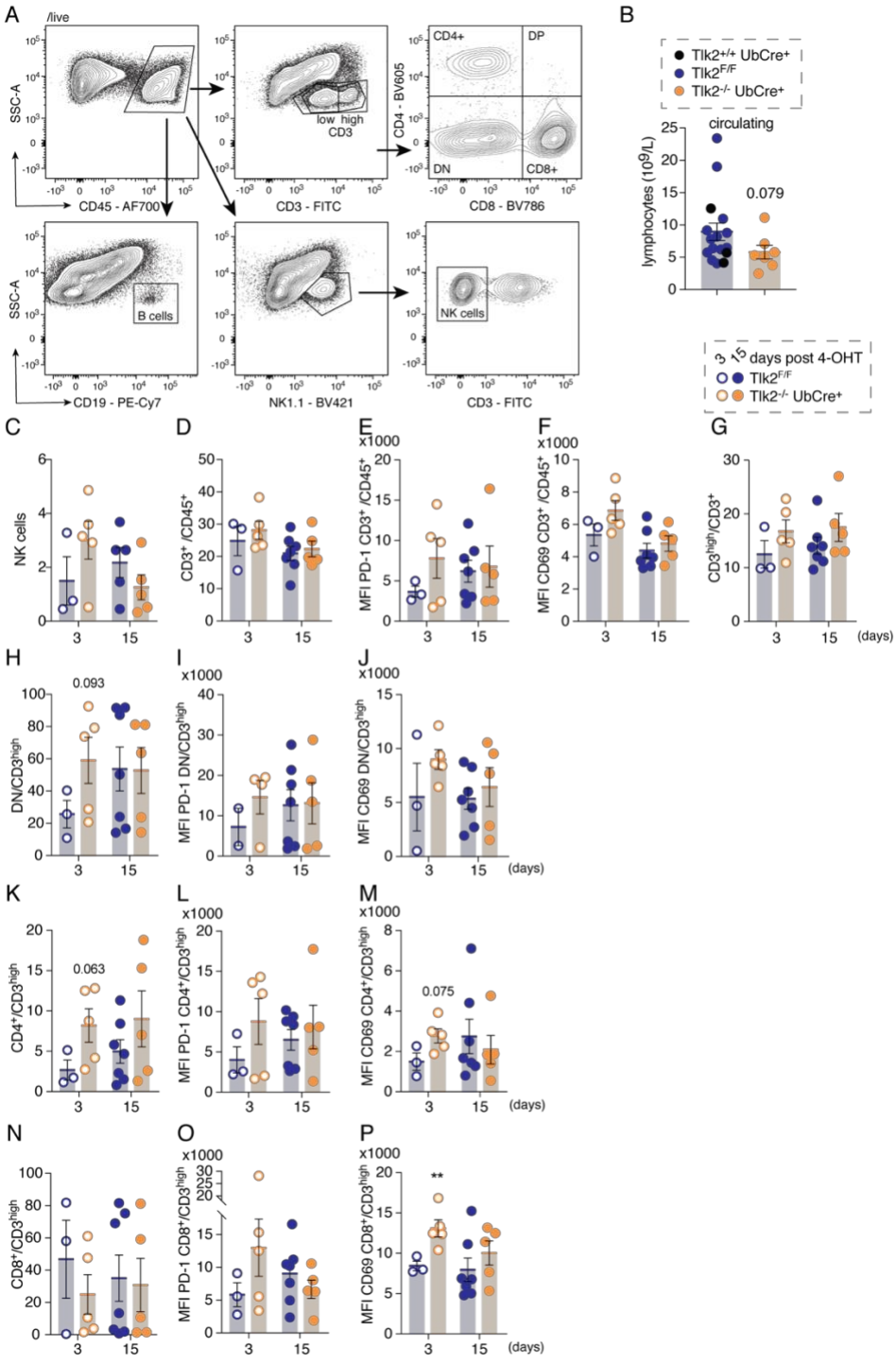

A
/live
SSC-A
CD45 - AF700
low high
CD3
CD3 - FITC
CD4+
DP
DN
CD8+
CD4 - BV605
CD8 - BV786
B cells
CD19 - PE-Cy7
NK cells
NK1.1 - BV421
CD3 - FITC
B
Tlk2+/+ UbCre+
Tlk2F/F
Tlk2-/- UbCre+
circulating
lymphocytes (10⁹/L)
0.079
3 15 days post 4-OHT
Tlk2F/F
Tlk2-/- UbCre+
C
NK cells
D
CD3+ /CD45+
E
x1000
MFI PD-1 CD3+ /CD45+
F
x1000
MFI CD69 CD3+ /CD45+
G
CD3high/CD3+
(days)
H
0.093
DN/CD3high
I
x1000
MFI PD-1 DN/CD3high
J
x1000
MFI CD69 DN/CD3high
(days)
K
0.063
CD4+/CD3high
L
x1000
MFI PD-1 CD4+/CD3high
M
x1000
0.075
MFI CD69 CD4+/CD3high
(days)
N
CD8+/CD3high
O
x1000
MFI PD-1 CD8+/CD3high
P
x1000
**
MFI CD69 CD8+/CD3high
(days)

Figure R7. Early increased lymphoid infiltration in TLK2 knockout tumors.

(**A**) Gating strategy used to study the lymphoid populations within the tumor: B cells ($CD45^+/CD3^-/CD19^+$), NK cells ($CD45^+/CD3^-/NK1.1^+$), T cells ($CD45^+/CD3^+$ with or without CD4 and CD8 expression). (**B**) Circulating lymphocytes (B+T cells) / liter (L) assessed by a complete blood count (CBC) performed at day 15 post 4-OHT treatment. **(C**, **D**, **G**, **H**, **K**, and **N**) percentage of the stated populations /$CD45^+$ cells. NK cells (**D**), T cells (**D**), activated T cells ($CD3^{high}$) (**G**), DN/$CD3^{high}$ (**H**), $CD4^+/CD3^{high}$ (**K**), and $CD8^+/CD3^{high}$ (**N**) at days 3 and 15 post 4-OHT treatment. (**E**, **I**, **L** and **O**) Mean Fluorescence Intensity (MFI) of PD-L1 of the indicated populations at day 15 post 4-OHT treatment. T cells (**E**), DN/$CD3^{high}$ (**I**), $CD4^+/CD3^{high}$ (**L**), and $CD8^+/CD3^{high}$ (**O**). (**F**, **J**, **M** and **P**) Mean Fluorescence Intensity (MFI) of CD69 of the indicated populations at day 15 post 4-OHT treatment. T cells (**F**), DN/$CD3^{high}$ (**J**), $CD4^+/CD3^{high}$ (**M**), and $CD8^+/CD3^{high}$ (**P**). Error bars represent SEM. Statistical significance was determined using an unpaired t test with Welch's correction (**P<0.01) and comparing the genotypes at the same day pot 4-OHT treatment. DN; double-negative ($CD4^-/CD8^-$), DP; double-positive ($CD4^+/CD8^+$).

Testing of small molecule inhibitors of TLK activity

TLK2 is overexpressed in several cancer types (Kim et al., 2016, Lee et al., 2018), and the fact that genetic ablation of TLK2 reduces tumor growth in a mouse model of breast cancer (Figure R2) suggests that TLK2 would be a good target for anti-cancer therapy. Potential TLK inhibitors have been identified after screening several compound libraries (Ronald et al., 2013) and publicly available kinase profiling data sets (Gao et al., 2013). Although these compounds inhibit TLKs *in vitro* at high concentrations (Kim et al., 2016, Mortuza et al., 2018) they lack specificity towards TLKs.

Our collaborators in Prof. Nathanael Gray's lab (Dana-Farber Cancer Institute, Boston, MA) provided 3 small molecule inhibitors, namely TLKi#1, TLKi#2, and TLKi#3, that they identified as potential TLK inhibitors in larger scale screenings of their in-house compounds. To test the small molecules' inhibitory activity *in vitro*, we analyzed substrate phosphorylation using TLK1 and TLK2 immunoprecipitated after overexpression in AD293 cells, and Myelin Basic Protein (MBP), a commonly used kinase substrate surrogate. We incubated the kinase with the inhibitors for 30 mins at room temperature prior to the addition of radioactive isotope Phosphorus-32 (^{32}P) (Figure R8A). We tested them

alongside 2 commercially available compounds that were reported to be TLK1 inhibitors (Ronald et al., 2013). Thioridazine (THD) and Fluphenazine (FF) are members of the family of phenothiazine antipsychotics, and THD was widely used for the treatment of schizophrenia until it was withdrawn in 2005 due to its severe side effects. THD was shown to inhibit TLK1 *in vitro* and reduce tumor size alone and in combination with low dose doxorubicin *in vivo*, while FF was shown not to inhibit TLK1 (Ronald et al., 2013). In addition, we used Staurosporine an ATP-competitive kinase inhibitor that we previously showed inhibits TLK activity *in vitro*, as a positive control (Mortuza et al., 2018). As expected, Staurosporine inhibits TLK1 and TLK2 catalytic activity, while FF did not (Figures R8B-E). Both TLKi#1 and TLKi#3 showed potent TLK1 and TLK2 inhibition on the substrate MBP and auto-phosphorylation (Figures R8B-E), which is needed for its activation (Mortuza et al., 2018). In our hands, THD did not inhibit TLK1 or TLK2 and we therefore did not pursue it further (Ronald et al., 2013).

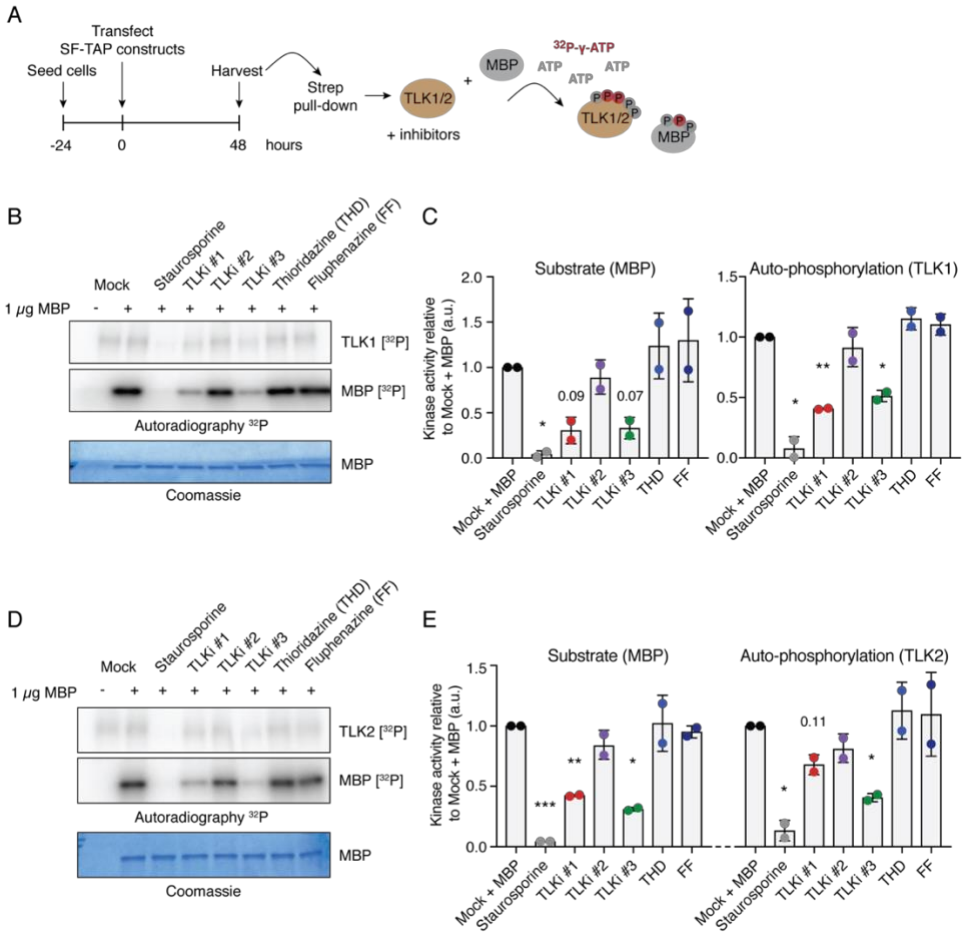


Figure R8. Testing of small molecule inhibitors of TLK activity.

(**A**) Schematic on the *in vitro* kinase assays used to assess the small molecules inhibitory capacity. TLK1 and TLK2 were overexpressed in AD293 cells and immunoprecipitated by Strep pull downs. The kinases were pre-incubated with the inhibitors during 30 mins in kinase buffer containing ATP at RT prior to addition of the ^{32}P. (**B**) *In vitro* kinase assay of TLK1 in the presence or absence of MBP and the indicated inhibitors. All inhibitors were used at 10 μM except for THD (16 μM) and FF (20 μM). The Coomassie staining shows equal loading of MBP. (**C**) Quantification of the TLK1 kinase assays shown in B. (**D**) *In vitro* kinase assay of TLK1 in the presence or absence of MBP and the indicated inhibitors. All inhibitors were used at 10 μM except for THD (16 μM) and FF (20 μM). The Coomassie staining shows equal loading of the MBP. (**E**) Quantification of the TLK1 kinase assays shown in D. Error bars represent SD of two biological replicates. ANOVA followed by Dunnett's test was applied when comparing multiple experimental groups against a control (Mock + MBP). Statistical significance was considered when adjusted p-values were ***P<0.001, **P<0.01, *P<0.05. All tests were

performed after log-transformation of the data. However, for clarity of representation, data are shown in the original scale. Statistical analysis was carried out with the multcomp (Hothorn et al., 2008) R package.

TLKi#1 and TLKi#3 synergize with ATR, CHK1, and PARP inhibition

We previously showed that TLK depletion with siRNAs sensitizes cells to cell cycle checkpoint inactivation and PARP inhibition (Lee et al., 2018). This synthetic lethal relationship could be therapeutically exploited to treat a broad range of cancers. TLKi#1 and TLKi#3 showed promising inhibition of TLKs *in vitro* (Figure R8). To test if these novel compounds have the same effects in cells in combination with checkpoint and PARP inhibitors, we set up an experiment to assess cell viability in a high-throughput manner. We used the AlamarBlue® (AB) cell viability assay. AB uses the capacity of living cells to continuously reduce resazurin (a non-toxic, non-fluorescent blue compound) into resorufin (a highly fluorescent red compound) as a quantitative readout for cellular proliferation (Figure R9A).

First, we tested the toxicity of TLK1i#1 and TLKi#3 in U-2-OS cells and they did not appear to be toxic at high concentrations (10 μM) (not shown). In order to observe a synthetic lethal interaction, we needed to identify the highest concentration of the compounds that does not alter the viability of the cells. To do so, we titrated down TLK1i#1 and TLKi#3, as well as the checkpoint inhibitors CHK1i (PFA477736) and ATRi (EP-46464), and the PARPi Olaparib (AZD2281) (not shown). We identified the following concentrations as optimal for this assay: TLK1i#1 (500 nM), TLKi#3 (20 nM), CHK1i (50 nM), ATRi (1 μM) and PARPi (500 nM). Next, we treated the cells with the inhibitors 24 hours after seeding, and performed the AB assay every 72 hours (Figure R9B). As with siRNA silencing, we observed synthetic lethality when combining the novel TLK inhibitors with checkpoint and PARP inhibitors (Figures R9C and R9D). The synergistic activity was calculated by a t-test of the combination of the inhibitors versus a Bliss independence baseline (Liu et al., 2018).

Although TLKi#1 and TLKi#3 do not specifically target TLKs, these results suggest that TLK kinases represent viable druggable targets for cancer therapy either alone or in combination with ATR, CHK1 or PARP inhibitors. Further

effort needs to be made in order to design more specific and potent TLK inhibitors that can be safely used in clinical trials in the future.

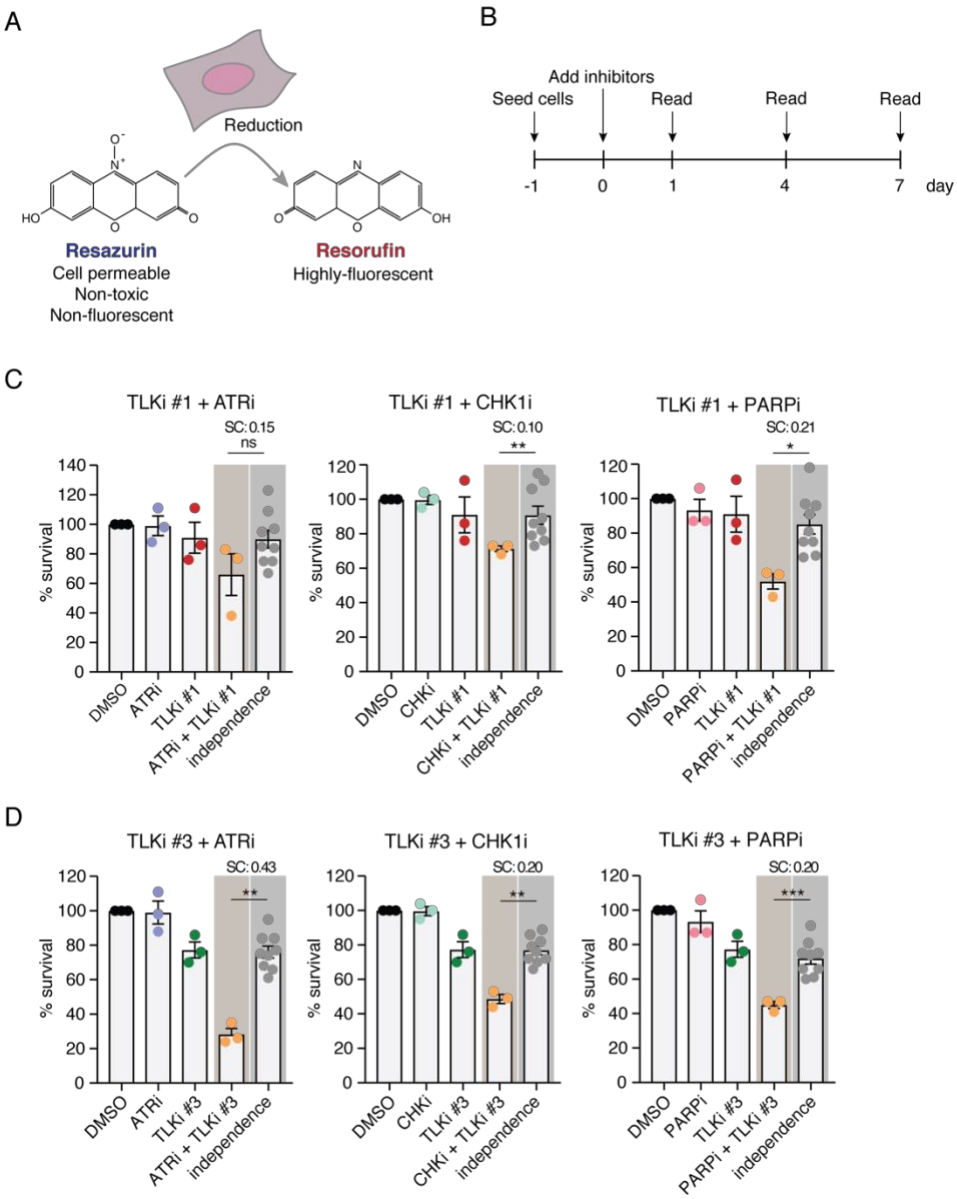


Figure R9. TLKi#1 and TLKi#3 synergise with ATR, CHK1, and PARP inhibition.

(**A**) AlamarBlue® (AB) cell viability assay. AB is a solution based in resazurin (a non-toxic, non-fluorescent blue compound) that is reduced by living cells into resorufin (a

highly fluorescent red compound). The fluorescence emitted by resorufin can be read with a fluorescent plate reader and be used as a quantitative readout for cellular proliferation. (**B**) Experimental layout. Cells were treated with the inhibitors 24 hours after seeding, and proliferation was read every 72 hours. (**C**) AB results of TLKi#1 in combination with ATRi, CHK1i, and PARPi. (**D**) AB results of TLKi#3 in combination with ATRi, CHK1i, and PARPi. The inhibitors were used at the following concentrations: TLK1i#1 (500 nM), TLKi#3 (20 nM), CHK1i (50 nM), ATRi (1 μM) and PARPi (500 nM). In (**C**) and (**D**) the "independence" column (grey background) shows a Bliss independence model of additive activity of the two treatments, against which the combined treatment (orange background) is tested to estimate synergistic activity (Liu et al., 2018). Error bars represent SEM of three biological replicates. Statistical significance was determined using a two-tailed t-test (***P<0.001, **P<0.01, *P<0.05). SC, synergy score.

Chapter 2

TLK2 loss leads to inflammation in the developing brain

Marina Villamor-Payà[1], María José Barallobre[2], Adrià Caballé Mestres[1], María Sanchiz-Calvo[1], Lluís Palenzuela[1], and Travis H. Stracker[1, 3].

[1]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona 08028, Spain

[2]Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER), Spain; Institut de Biologia Molecular de Barcelona (IBMB), 08028 Barcelona, Spain

[3]Radiation Oncology Branch, National Cancer Institute, Bethesda, MD 20892, USA

Statement of contribution:

M Villamor-Payà designed, performed and analyzed all experiments.

MJ Barallobre contributed to experimental design (RNAseq and IHC experiments) and helped M Villamor-Payà with the necropsies to obtain brain and telencephalon from embryos.

A Caballé Mestres performed all bioinformatic analysis, with constant feedback from M Villamor-Payà.

M Sanchiz-Calvo helped with the first analysis of the RNAseq that led to Figure R15 and co-performed the qPCRs in Figure R13C-D with M Villamor-Payà.

Lluís Palenzuela helped maintain mouse cohorts and with mouse genotyping.

TH Stracker supervised the whole project.

TLK2 was recently identified as a risk gene for intellectual disability (ID) (Lelieveld et al., 2016), and point mutations in TLK2 were later associated with a distinct neurodevelopmental disorder - Mental Retardation Autosomal Dominant (MRD57, MIM: 618050) (Reijnders et al., 2018). MRD57 is clinically characterized by autism-spectrum disorder (ASD), ID, behavioral problems, growth delay and facial dysmorphism. A subset of cases also presents with gastrointestinal problems, seizures, skeletal malformations and ocular problems.

With all the knowledge we have gathered about TLKs over the years, we came up with several non-exclusive hypotheses that could explain why mutations that affect TLK2 activity are involved in neurodevelopmental disorders (NDDs); 1) TLK2 loss compromises placental development in knockout mice and could be triggering an inflammatory response in the placenta. 2) TLK2 was identified as a modifier of micro-exon inclusion *in vitro*, and could be causing defects in brain-specific micro-exon splicing required for neural development. 3) TLK2 defects could impair DNA replication and cause neuroprogrenitor attrition due to RS. All of these scenarios have been associated with ASD and other NDDs and are consistent with existing data (Carpentier et al., 2011, Irimia et al., 2014, Choi et al., 2016, Straughen et al., 2017, Kim et al., 2017, Gonatopoulos-Pournatzis et al., 2018, Segura-Bayona et al., 2020). We therefore decided to address these hypotheses *in vivo* using our TLK2 mouse model.

TLK2 knockout placentas do not exhibit increased inflammation

The generation of mice deficient for *Tlk1* or *Tlk2* with a gene-trap system showed that *Tlk2* (but not *Tlk1*) was required for normal placental development and its loss led to embryonic lethality. Extensive characterization of the placental phenotype was performed previously (Segura-Bayona et al., 2017). Briefly, $Tlk2^{T/T}$ placentas were thinner, poorly vascularized and composed of less differentiated and more immature trophoblast cells. Surpassing the placental phenotype with a Sox2-Cre system allowed for the generation of $Tlk2^{-/-}$ adult mice, that present with a mild developmental delay but not any other overt phenotypes.

Acute and chronic placental inflammation have been associated with increased risk of ASD (Straughen et al., 2017), as well as maternal viral and bacterial

infections during pregnancy (Patterson, 2011, Zerbo et al., 2015, Kim et al., 2017). In fact, maternal immune activation (MIA) by injection of the viral mimic poly(I:C) during pregnancy, is a widely used mouse model for autism (Malkova et al., 2012), where the male offspring present with classical autistic behaviors.

The fact that TLK2 is essential for proper placental development (Segura-Bayona et al., 2017) and that TLK depletion *in vitro* generates a potent antiviral and innate immune response (Segura-Bayona et al., 2020), made us wonder if placental inflammation could be the cause of ASD in patients with *TLK2* mutations. To assess if *Tlk2* knockout placentas exhibited increased inflammation, we performed RNA sequencing (RNA-seq) of $Tlk2^{+/+}$, $Tlk2^{+/-}$ and $Tlk2^{-/-}$ male placentas at embryonic day E13.5 (Figure R10A). At this embryonic age, $Tlk2^{-/-}$ embryos are smaller but have not yet perished. Genotyping of the embryos identified $Tlk2^{+}$ and $Tlk2^{-}$ alleles (Figure R10B). Reads per kilobase of transcript, per million mapped reads (RPKM) shows no counts of exon 4 in the RNA-seq data (Figure R10C) and reduced levels of TLK2 protein were detected in $Tlk2^{-/-}$ placental tissue (Figure R10D), proving that the system worked well.

Differential expression between genotype condition pairs was performed using the Limma R package (Ritchie et al., 2015) on the gene level intensities. This analysis shows that the highest upregulated genes in the $Tlk2^{-/-}$ samples are placental hormones, including members of the prolactin family and placental growth factor (PGF) (Figure R10E). We also performed gene set enrichment analysis (GSEA) on Hallmark terms (Liberzon et al., 2015), Gene Ontology (GO) terms (Mi et al., 2019), and KEGG terms (Kanehisa and Goto, 2000). The most striking result is that heparan sulphate biosynthesis is strongly downregulated in the $Tlk2^{-/-}$ placentas (Figure R10G). Interestingly, both exposure to elevated levels of prolactin *in utero* and heparan sulphate deficiencies have been linked to ASD (Yarlagadda et al., 2015, Perez et al., 2016).

That we could not readily observe an increased inflammatory response in the $Tlk2^{-/-}$ placentas (Figure R10F) suggests that placental inflammation is most likely not a contributing cause of ASD in patients with *TLK2* mutations. However, alterations in pathways that contribute to ASD risk, such as the heparan sulphate biosynthesis, were altered in the $Tlk2^{-/-}$ placentas, so we cannot rule out a potential placental contribution to the human disorder.

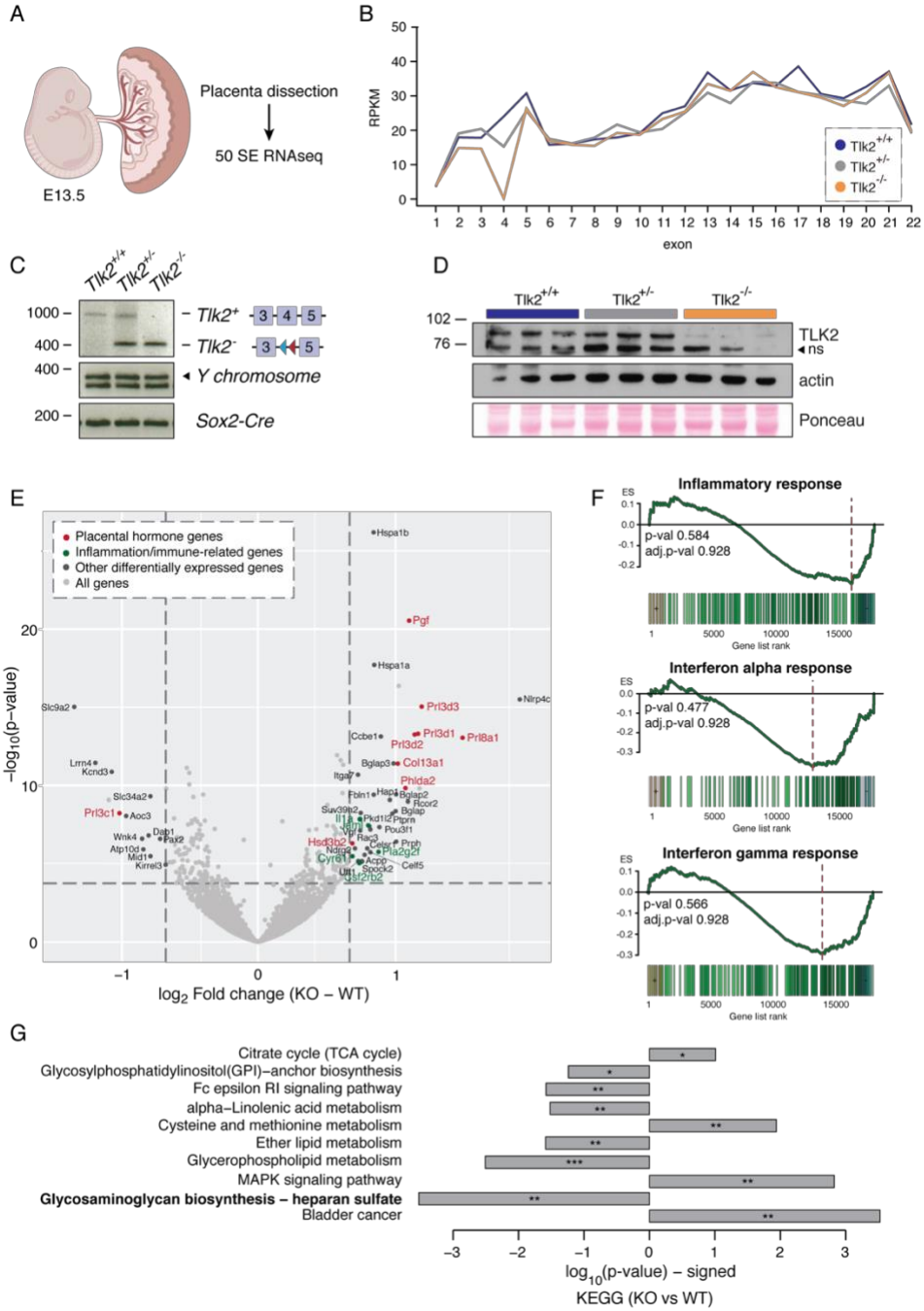


Figure R10. TLK2 knockout placentas do not exhibit increased inflammation.

(**A**) Experimental layout. 3 E13.5 male placentas per genotype (*Tlk2*$^{+/+}$, *Tlk2*$^{+/-}$ and *Tlk2*$^{-/-}$) were sequenced SE 50 bp. (**B**) Reads per kilobase of transcript, per million mapped

reads (RPKM) show *Tlk2* at the exon level and confirm that exon 4 is knocked out in *Tlk2*$^{-/-}$ placentas. (**C**) Example of the PCR-based genotyping of *Tlk2*$^{+/+}$, *Tlk2*$^{+/-}$ and *Tlk2*$^{-/-}$ embryos. (**D**) Western blotting of TLK2 protein levels of the indicated genotypes from placental tissue dissections. Actin and Ponceau staining show equal loading. (**E**) Volcano plot representing the gene expression profile of *Tlk2*$^{-/-}$ (KO) versus *Tlk2*$^{+/+}$ (WT) obtained from RNA-seq of E13.5 embryonic placentas (n = 3 samples/genotype). Light gray dots indicate genes, dark gray indicate differentially expressed genes. Red dots indicate genes encoding for placental hormones while green dots indicate genes directly or indirectly linked to inflammation/immune system. Dotted lines indicate threshold. Differential expression between genotype condition pairs (KO vs WT) was performed using the Limma R package (Ritchie et al., 2015) on the gene level intensities. (**F**) Gene set enrichment analysis (GSEA) where genes were annotated to Hallmark terms (Liberzon et al., 2015). Inflammatory, interferon alpha and gamma responses are shown. (**G**) GSEA of KEGG terms (Kanehisa and Goto, 2000) of RNA-seq differentially up and down-regulated genes in *Tlk2*$^{-/-}$ (KO) placentas as compared to *Tlk2*$^{+/+}$ (WT). Only the top significant hits are shown. ns, non-specific.

TLK2 knockout telencephalons have normal alternative splicing

Our next hypothesis was that TLK2 loss could be mis-regulating micro-exon splicing in the brain of *Tlk2*$^{-/-}$ mice. TLK2 was recently identified in a genome-wide CRISPR-based screen as one of ~200 novel regulators of neuronal micro-exon splicing, as its depletion identified it as a positive regulator of nSR100/SRRM4 gene expression (Gonatopoulos-Pournatzis et al., 2018). Micro-exons are very short exons (3 to 27 nucleotides) that are highly conserved through evolution and frequently misregulated in the brains of autistic individuals (Irimia et al., 2014, Quesnel-Vallieres et al., 2015). Most neural micro-exons are regulated by the neuronal-specific Ser/Arg repeat-related protein of 100 kDa (nSR100/SRRM4), and mice haploinsufficient for nSR100/SRRM4 display hallmarks of ASD (Quesnel-Vallieres et al., 2016).

To study micro-exon usage and alternative splicing we performed RNA-seq on telencephalon tissue of E18.5 embryos. We sequenced PE 125 bp at a depth of 70 million reads/sample to be able to capture the micro-exon splicing events (Figure R11A-D). We used Vasttools (Tapial et al., 2017) for alignment and alternative splicing quantification. All observations from the same genotype were combined for percentage of spliced-in (PSI) calculations and events with less than

50 reads in both *Tlk2*$^{+/+}$ Sox2-Cre^{+} and *Tlk2*$^{-/-}$ Sox2-Cre^{+} were filtered out. We distinguished between exon skipping in micro-exons (nt ≤ 27) (Figure R12A) and alternative-exons (nt > 27) (Figure R12B). In these graphs, each dot represents the percentage of times a micro-exon (Figure R12A) or an alternative exon (Figure R12B) is included (spliced-in) in average in the transcript (average PSI). The boxplots represent the differential PSI (ΔPSI) between *Tlk2*$^{+/+}$ Sox2-Cre^{+} and *Tlk2*$^{-/-}$ Sox2-Cre^{+}. Although there are some micro-exons differentially spliced between *Tlk2*$^{+/+}$ Sox2-Cre^{+} (WT) and *Tlk2*$^{-/-}$ Sox2-Cre^{+} (KO), it was not statistically significant (Figure R12A). In the case of all alternative splicing events (Figure R12B), no changes were observed between genotypes. It is important to note that our definition of differential splicing events was quite relaxed when compared to other studies (Quesnel-Vallieres et al., 2016). As micro-exons are significantly enriched in genes that have been genetically linked to ASD (Irimia et al., 2014), we decided to compare our list of genes with mis-regulated splicing to the Simon's Foundation Autism Research Initiative (SFARI) genesets (Banerjee-Basu and Packer, 2010, Abrahams et al., 2013) and see if there was an enrichment in ASD-susceptibility genes. There was not a statistically significant overlap in micro-exons (Figure R12C) or all alternative splicing events (Figure R12D), meaning that the number of genes that overlapped between our geneset and SFARI could as well be random.

This extensive splicing analysis together with the fact that the expression levels of the neuronal micro-exon splicing effector nSR100/SRRM4 remains intact in the *Tlk2*$^{-/-}$ Sox2-Cre^{+} telencephalons (Figure R13D) suggested that mis-regulation of micro-exon splicing likely does not significantly contribute to ASD in patients with *TLK2* point mutations. However, an important consideration must be considered. We performed bulk RNA-seq, so all the cell types of the telencephalon were included, but micro-exon splicing happens predominantly in neurons. If there was a shift in the cell populations percentages, as for example a big decrease in neurons or an increase in another cell population, the differences in micro-exon splicing could be diluted.

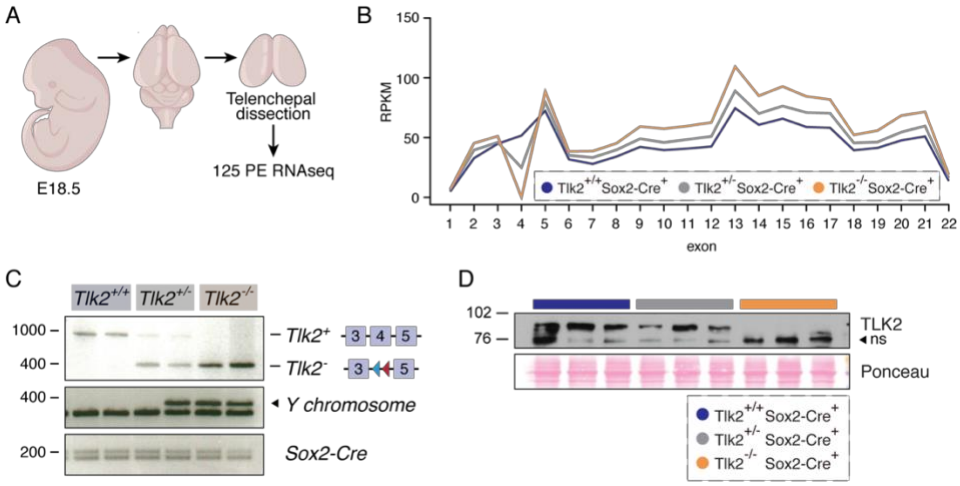


Figure R11. Experimental workflow for telencephalon RNA-seq.

(**A**) 6 E18.5 embryos per genotype (*Tlk2*$^{+/+}$ Sox2-Cre^{+}, *Tlk2*$^{+/-}$ Sox2-Cre^{+} and *Tlk2*$^{-/-}$ Sox2-Cre^{+}, 3 males and 3 females) were used for telencephalon isolation and posterior paired-end RNA-seq at a length of 125 bp. (**B**) Reads per kilobase of transcript, per million mapped reads (RPKM) show TLK2 at the exon level. Exon 4 is not present on the *Tlk2*$^{-/-}$ Sox2-Cre^{+} telencephalons. (**C**) Example of the PCR-based genotyping of *Tlk2*$^{+/+}$, *Tlk2*$^{+/-}$ and *Tlk2*$^{-/-}$ embryos, Sox2-Cre and sex PCR. (**D**) Western blotting of TLK2 protein levels of the indicated genotypes from telencephalon tissue dissections. Ponceau staining shows equal loading. ns, non-specific.

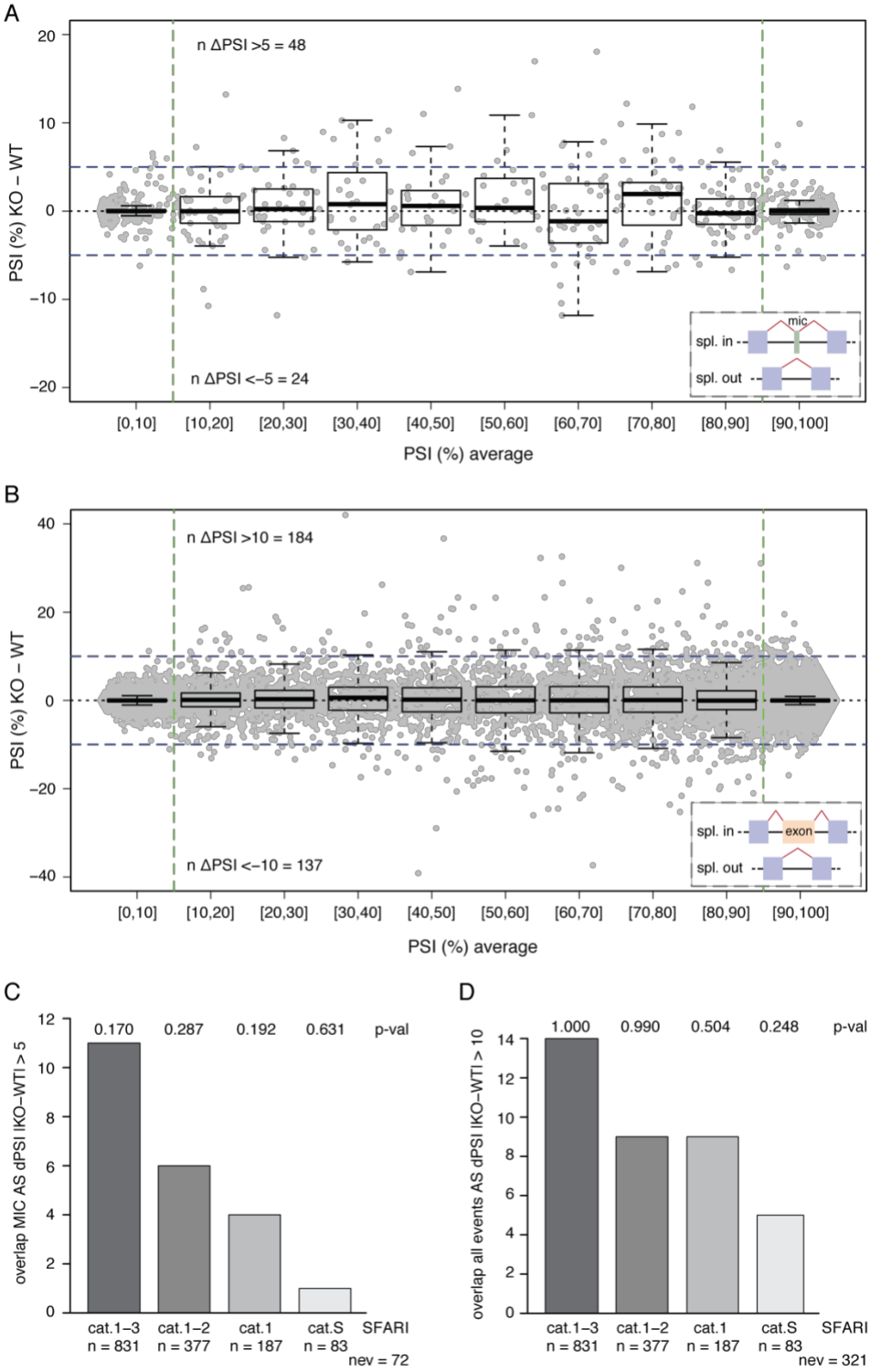

A
n ΔPSI >5 = 48
n ΔPSI <–5 = 24
PSI (%) KO – WT
PSI (%) average
[0,10]
[10,20]
[20,30]
[30,40]
[40,50]
[50,60]
[60,70]
[70,80]
[80,90]
[90,100]
mic
spl. in
spl. out
B
n ΔPSI >10 = 184
n ΔPSI <–10 = 137
exon
C
overlap MIC AS dPSI IKO–WTI > 5
0.170
0.287
0.192
0.631
p-val
cat.1–3 n = 831
cat.1–2 n = 377
cat.1 n = 187
cat.S n = 83
SFARI nev = 72
D
overlap all events AS dPSI IKO–WTI > 10
1.000
0.990
0.504
0.248
SFARI nev = 321

Figure R12. TLK2 knockout telencephalons have normal alternative splicing.

(**A** and **B**) Micro-exons (nt ≤ 27) (**A**) and alternative exons (nt > 27) (**B**) splicing analysis. Vasttools (Tapial et al., 2017) version 2.2.2. was used for alignment and alternative splicing quantification. All observations from the same genotype were combined for percentage of spliced-in (PSI) calculations. Only events with a minimum of 50 reads in both *Tlk2*$^{+/+}$ Sox2-Cre^{+} and *Tlk2*$^{-/-}$ Sox2-Cre^{+} samples were considered. Differential alternative splicing events were defined by a |ΔPSI| > 5 in micro-exons and by a |ΔPSI| > 10 in alternative-exons (represented by blue dotted lines). Green dotted lines represent Exon skipping (ExS) events with 10≤ PSI ≤ 90. (**C** and **D**) Enrichment analysis results using SFARI genesets (gene.sfari.org) of splicing events of micro-exons with a |ΔPSI| > 5 (**C**) and alternative-exons with a |ΔPSI| > 10 (**D**). SFARI categories 1 (high confidence, gene listed in (Satterstrom et al., 2020)), 2 (strong candidate), 3 (suggested evidence) and S (syndromic). Spl, spliced; AS, alternative splicing; MIC, micro-exon.

TLK2 knockout embryos exhibit elevated transcription of inflammatory genes

We next analyzed differential RNA expression between genotype condition pairs of the E18.5 telencephalons. Principal component analysis (PCA) showed that sex and genotype account for a large percentage of the differences observed between samples. Sex (PCA1) explains 7.6% of the variability, while genotype (PCA2) 7.0%. *Tlk2*$^{-/-}$ Sox2-Cre^{+} (KO) and *Tlk2*$^{+/+}$ Sox2-Cre^{+} (WT) are the most different with *Tlk2*$^{+/-}$ Sox2-Cre^{+} (HET) present an intermediate pattern with half the samples overlapping with the WT (Figure R13A). The most highly enriched GO categories involved innate immune response and interferon genes, whilst the strongest downregulation was in histone variant genes (Figure R13B). We further validated these findings using quantitative real-time PCR (Figure R13C). These results are consistent with our previous results that TLK loss *in vitro* leads to reduced histone density and an effective immune response (Segura-Bayona et al., 2020). It is also important to note that within the upregulated genes there was an over representation of genes predominantly expressed by glial and vascular cells (Zhang et al., 2014) (Figure R13B).

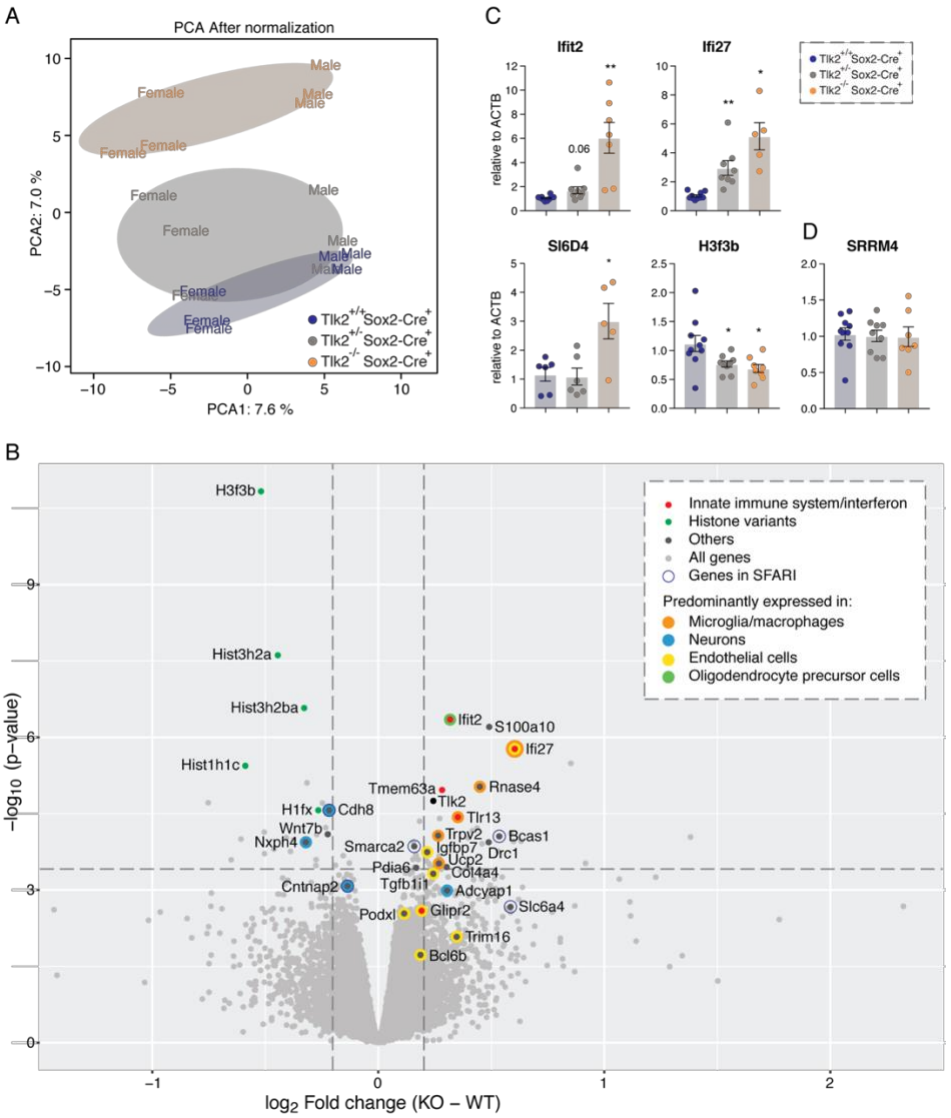


Figure R13. Innate immune response genes and histone variants are differentially regulated in TLK2 knockout telencephalons.

Gene expression results from RNA-seq of E18.5 embryonic telencephalons (n = 6 samples/genotype). (**A**) Principal component analysis (PCA) shows how the main components that drive the differences between our RNA-seq samples are genotype and sex. (**B**) Volcano plot representing the gene expression profile of *Tlk2*$^{-/-}$ Sox2-Cre^{+} (KO) versus *Tlk2*$^{+/+}$ Sox2-Cre^{+} (WT). Light gray dots indicate genes, dark gray indicate genes of interest, green dots represent histone variants (downregulated) whereas red dots represent genes belonging to the innate immune system / Interferon response (upregulated). Blue circles mark genes present in the SFARI database (gene.sfari.org).

Genes that are predominantly expressed in a specific brain cell type are shown (brainrnaseq.org; (Zhang et al., 2014)). Dotted lines indicate threshold. Differential expression between genotype condition pairs (KO vs WT) was performed using the Limma R package (Ritchie et al., 2015) on the gene level intensities. (**C** and **D**) Quantitative real time PCR analysis of some of the most upregulated (*Ifit2*, *Ifi27*, *Sl6D4*) and downregulated (*H3hf3b*) genes in the *Tlk2*$^{-/-}$ Sox2-Cre^{+} telencephalons (**C**) and *nSR100/SRRM4* (**D**) from the RNA-seq shown in (**B**)**.** Each dot represents the mean of technical triplicates from a mouse of a given genotype. 5-10 mice / group were used. Data were normalized to unchanging expression levels of a housekeeping gene (ACTB) and to the mean of all the *Tlk2*$^{+/+}$ Sox2-Cre^{+}. Mean ± SEM is shown. An unpaired t-test with Welch's correction was used for statistical analysis (**P<0.01, *P<0.05).

GSEA confirmed that among categories up-regulated in the *Tlk2*$^{-/-}$ Sox2-Cre^{+} telencephalons there was inflammation, as well as IFN alpha and gamma response genes (Figure 14A). In addition, gene signatures of TNFα signaling via nuclear factor kB (NF-kB), apoptosis, hypoxia, and MTORC1 signaling were also up-regulated (Figure 14A). GO analysis showed in the Cell Component analysis how categories related to histones, nucleosome assembly and euchromatin were downregulated (Figure 14B).

Neuroinflammation is a prominent hallmark of many different neurodevelopmental disorders (Glass et al., 2010) and a growing number of studies support its involvement in ASD etiology (Pardo et al., 2005, Zimmerman et al., 2005, Liao et al., 2020). Increased mTOR-mediated signaling has been correlated with severity in idiopathic ASD (Sato, 2016, Ganesan et al., 2019), while pre-natal hypoxia (Burstyn et al., 2011, Driscoll et al., 2018), elevated IFN-γ (Warre-Cornish et al., 2020) and TNFα levels (Xie et al., 2017), and aberrant NF-kB expression have been shown to play a critical role in the neuroinflammation process in ASD pathogenesis (Young et al., 2011). Taken together, these data suggest that the telencephalons of our *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice present with several altered pathways corresponding to those implicated in ASD.

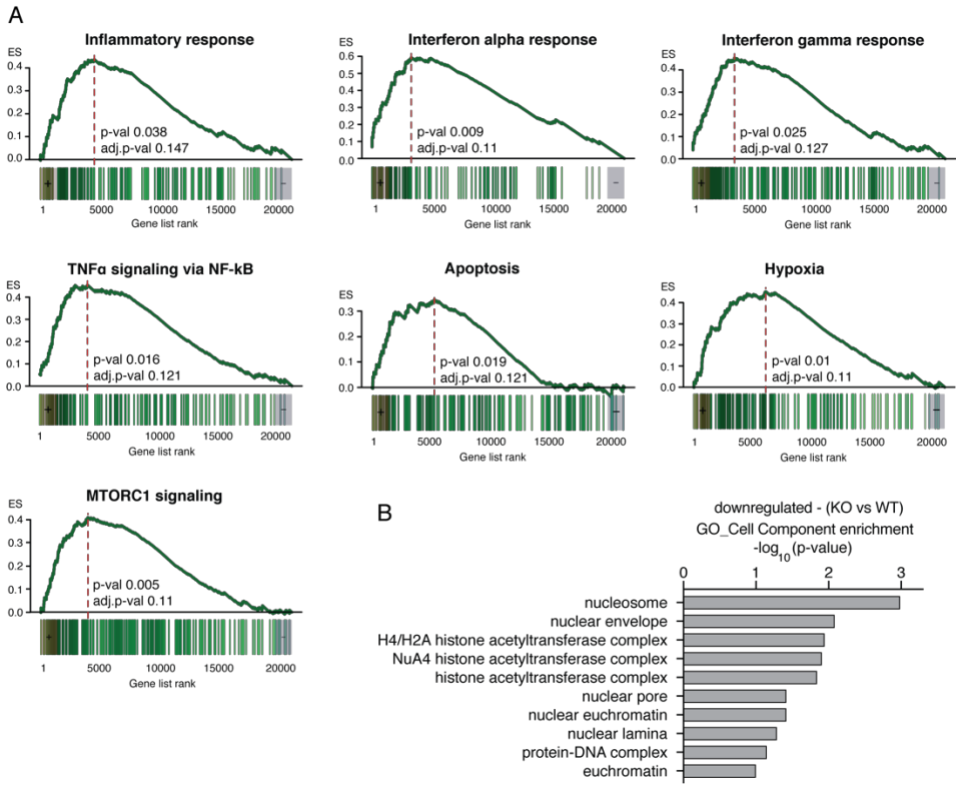


Figure R14. Increased expression of neuroinflammatory genes in TLK2 knockout embryos.

(**A**) GSEA where genes were annotated to Hallmark terms (Liberzon et al., 2015). Some of the categories upregulated in *Tlk2*$^{-/-}$ Sox2-Cre^{+} telencephalons are shown. (**B**) GO analysis (Cell Component enrichment) of RNA-seq differentially downregulated genes in *Tlk2*$^{-/-}$ Sox2-Cre^{+} (KO) telencephalons.

Stromal cell components are upregulated in the TLK2 knockout telencephalons

Neuroinflammation is generally considered a process in which microglia are activated for extended periods of time, therefore producing an excess of pro-inflammatory cytokines (Liao et al., 2020). An extensive review of postmortem studies of neuroinflammation in ASD (Liao et al., 2020) found that several studies identified alterations in cell number and activation markers of both astrocytes and microglia (Pardo et al., 2005, Vargas et al., 2005, Young et al., 2011), as well as abnormally high levels of proinflammatory cytokines and

chemokines that can be detected in the cerebrospinal fluid (CSF) of the patients (Krakowiak et al., 2017, Napolioni et al., 2013).

During our preliminary analysis we realized that most of our up-regulated genes corresponding to innate immunity and inflammation were expressed almost exclusively by microglial cells (Figure R13B) (Zhang et al., 2014). We decided to expand this analysis by comparing our RNA-seq results to the gene signatures defined by the Mouse Cell Atlas top 100 markers of every E18-brain cell type (Han et al., 2018). We observed a significant decrease in the gene expression signature of cortical neurons and an increase of several glial and stromal cell categories, such as Aldoc^{high} astrocytes, endothelial cells, macrophage/microglia and to a lesser extend oligodendrocytes; as well as some subsets of RGCs. A preliminary immunohistochemical analysis of E18.5 cortex showed an apparent increase in microglial cells in the cortex of *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice (Figure R16), as assessed by staining with Iba-1, a microglia-specific marker (Ito et al., 1998). While further immunohistochemical characterization on other cell populations, proliferation and DNA damage status are underway, we believe it is possible that the *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice have a premature onset of gliogenesis. That would explain the reduction in cortical neurons and the increase in astrocyte and RGC populations with overlapping gene expression (Supplementary table 4). These data suggest that the neuroinflammation we observed in the *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice are likely due to an increase in microglia and astrocytes, or their activity, which has been widely associated with NDD pathogenesis (Gandal et al., 2018b).

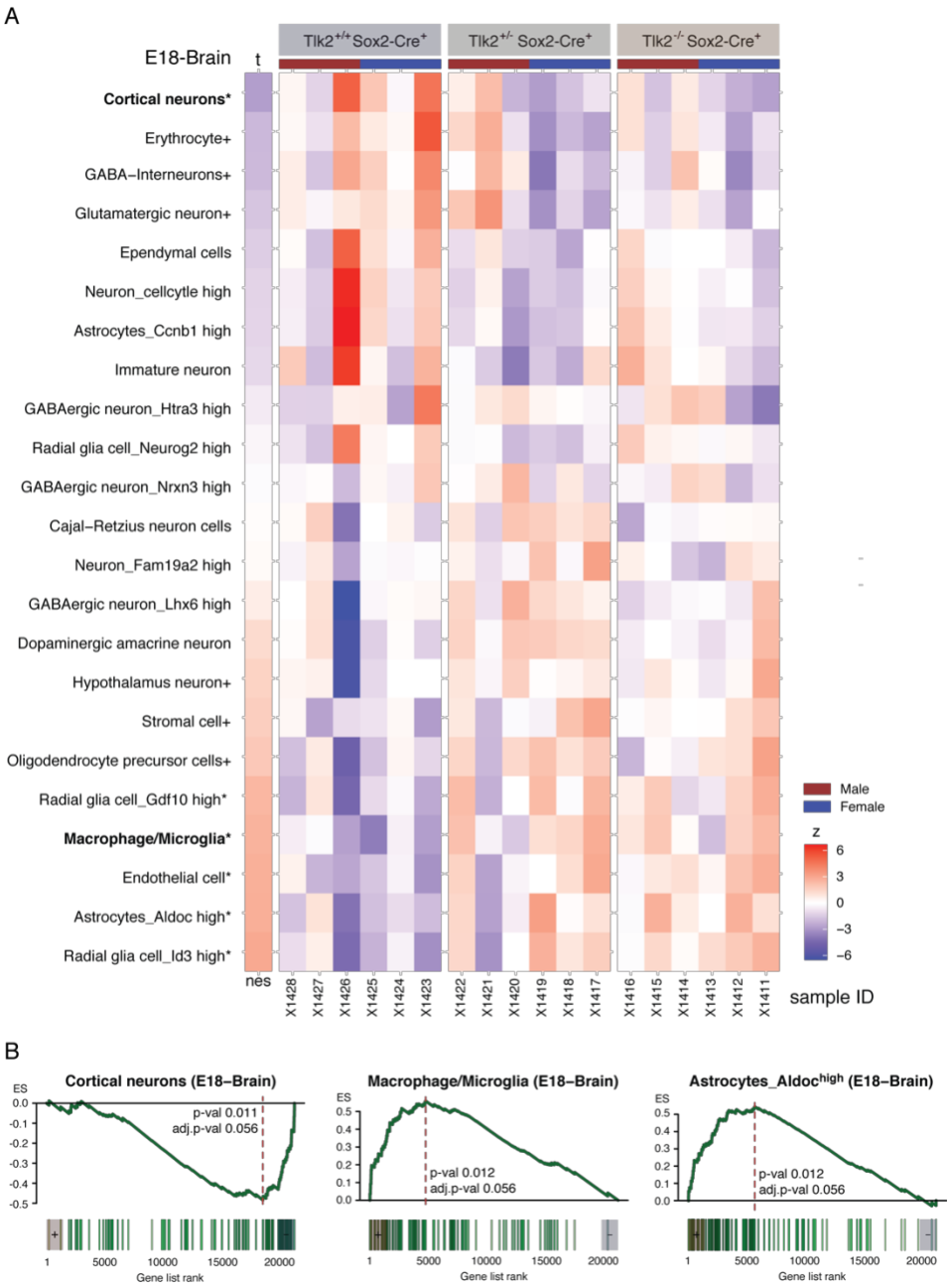


Figure R15. Stromal cell components are upregulated in the TLK2 knockout telencephalons.

(**A**) GSEA was performed employing the Casper gene level intensities and gene signatures defined by the Mouse Cell Atlas top 100 markers of every E18-brain cell type (Han et al., 2018), version 1.1. (*P<0.05, +P<0.25). Signatures with names finishing in

_gene name$^{low/high}$ were populations not described previously. (**B**) GSEA for the cortical neurons, microglia/macrophage, Aldoc^{high} astrocytes and signatures is shown.

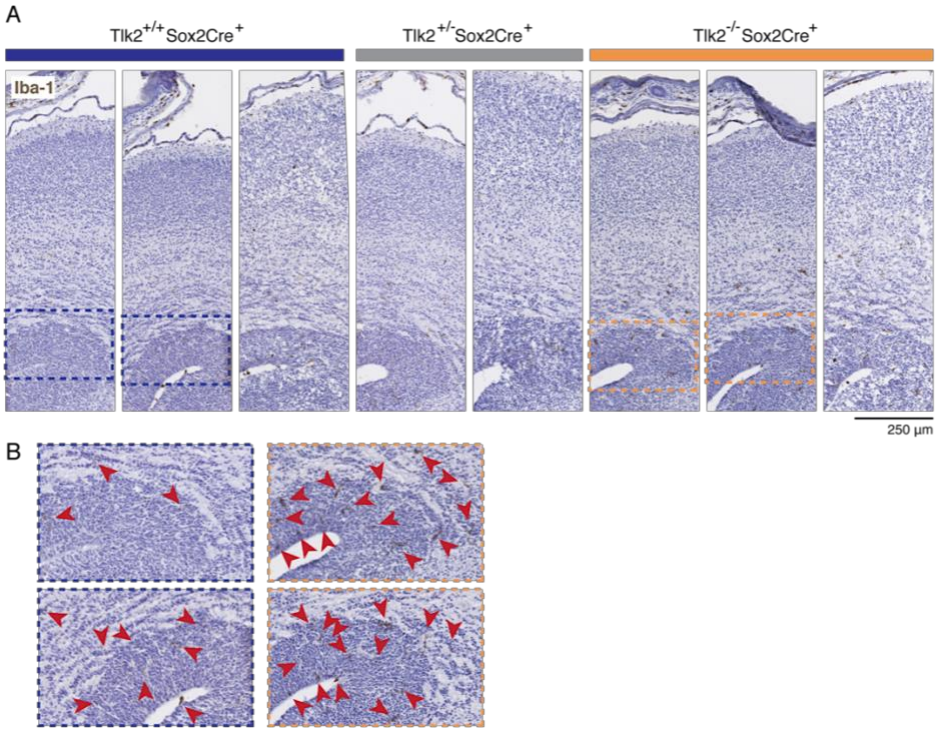


Figure R16. Increased number of microglial cells in the cortex of *Tlk2*$^{-/-}$ Sox2-Cre^{+} mice.

(**A**-**B**) Representative images of Iba-1 immunohistochemical staining of coronal sections of E18.5 embryonic heads of the indicated genotypes. Dashed square regions in (**A**) are zoomed in in (**B**) and Iba-1^{+} cells are indicated with red arrowheads.

Chapter 3

Functional analysis and interactome characterization of TLK2 variants identified in MRD57 patients

Marina Villamor-Payà[1], Maria Sanchiz-Calvo[1], Lisa Pavinato[2], Marina Gay[1], Marta Vilaseca[1], Gianluca Arauz-Garofalo[1], Cristina Andreoli[3], Valentina Prota[3], Elisa Giorgio[2], Andrea Ciolfi[4], Alessandro Bruselles[5, 6], Tommaso Pippucci[7], Francesca Clementina Radio[4], Paola Dimartino[8], Marco Tartaglia[4], Alfredo Brusco [2, 9], Travis H. Stracker [1, 10]

[1]Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), 08028 Barcelona, Spain

[2]Department of Medical Sciences, University of Turin, Turin, Italy

[3]Department of Environment and Health, Istituto Superiore di Sanità, Roma, Italy

[4]Genetics and Rare Diseases Research Division, Ospedale Pediatrico Bambino Gesu, Roma, Italy

[5]Department of Oncology and Molecular Medicine, Istituto Superiore di Sanita, Rome, Italy

[6]Istituto Superiore di Sanita, Roma, Italy

[7]Medical Genetics Unity, University of Bologna Hospital of Bologna Sant'Orsola-Malpighi Polyclinic, Bologna, Italy

[8]Department of Pediatrics and Public Health and Pediatric Sciences, University of Turin, Turin, Italy

[9]Unit of Medical Genetics, Azienda Ospedaliero Universitaria Città della Salute e della Scienza di Torino, Turin, Italy

[10]Radiation Oncology Branch, National Cancer Institute, Bethesda, MD 20892, USA

Statement of contribution:

M Villamor-Payà cloned the TLK2 mutants and designed and performed most experiments shown in Figures R20 to R27

M Sanchiz-Calvo co-prepared BioID-MS samples with M Villamor-Payà, performed localization experiments shown in Figure R21 and started experiments that led to Figure R26

L Pavinato collected the cases shown in Figure R17, performed variant confirmation and interpreted exome data

M Gay, M Vilaseca and G Arauz-Garofalo analysed BioID-MS data shown in Figures R22D and R23-R25

C Andreoli and V Prota performed the SCGE assays shown in Figure R18

E Giorgio interpreted exome data (Figure R17)

A Ciolfi, A Bruselles, T Pippucci, FC Radio, P Dimartino and M Tartaglia processed and analyzed WES data (Figure R17)

A Brusco designed and supervised the experiments shown R17 and R18

TH Stracker designed and supervised the experiments shown in R19-R27

Identification of 6 individuals with novel TLK2 variants

Our collaborators in Dr. Alfredo Brusco's lab re-evaluated the genomic information available for a large cohort of patients affected by ID and/or ASD who had previously been analyzed by whole exome sequencing (WES) and array-based comparative genomic hybridization (array-CGH). This included an in-house cohort (2,250 total samples, 736 affected) and the publicly available cohort from the Autism Sequencing Consortium (ASC, 35,584 total samples, 11,986 affected) (Satterstrom et al., 2020). We also performed a search in the DECIPHER database (Firth et al., 2009) for novel *TLK2* deletions. We identified six cases in three independent families with potentially harmful variants in the *TLK2* gene (Figure R17).

All patients had heterozygous *TLK2* mutations. Patients from families 1 and 2 carried point mutations in the Protein Kinase Domain (PKD) (Figure R17A), while the patient from family 3 carried a deletion in chromosome 17 encompassing *TLK2* (Figure R17F). In family 1 (Figures R17B and R17D), the affected child carried a *de novo* missense variant c.1652A>G; p.(Asp551Gly) not reported in The Genome Aggregation Database (gnomAD) and it was predicted to be damaging by multiple *in silico* predictors (Figure R19D). This variant localizes in a highly Constrained Coding Region (CCR) (>93th percentile) within the PKD (Havrilla et al., 2019). The patient was enrolled to ASC after being diagnosed with ASD.

In family 2, three affected siblings carried a premature stop variant (c.1423G>T; p.(Glu475Ter)) inherited from their affected mother (Figures R17C and R17E). This variant was not reported in gnomAD and was classified as pathogenic using the American College of Medical Genetics (ACMG) criteria (Reijnders et al., 2018, Topf et al., 2020). Patients 3 and 4 and the parents were subjected to trio WES, meanwhile patient 5 was born. The mother was diagnosed with mild ID after the (c.1423G>T; p.(Glu475Ter)) mutation was identified. All affected members from family 2 had microcephaly.

The case in family 3 was identified in the DECIPHER (Firth et al., 2009) database. The female carried a *de novo* deletion in 17q23.2 that was discovered by array-CGH. The minimum size of the deletion was 39 kb and the maximum size

87 kb (chr17-60683462-60722398) and included the *TLK2* gene (Figure R17F and R17G). The other genes affected by this deletion, *MRC2* and potentially *MARCH10*, are not predicted to be haploinsufficient (based on gnomAD), nor there is supporting literature implicating them in NDDs. We therefore believe they are unlikely to contribute to the observed pathology.

The patients (3 females and 3 males) were all of Caucasian ethnicity and ranged from 3 to 47 years of age. They presented with a broad range of behavioural disorders, including ASD, attention deficit hyperactivity disorder (ADHD), anxiety, short attention span and obsessive-compulsive behaviour (OCD). 4 patients had ID in the borderline (Intelligence quotient (IQ) 70-85) or low (IQ≤70) ranges. All the patients also presented with several dysmorphic facial features, and minor skeletal anomalies of the hands and feet were reported in 4 patients. Interestingly, some of the features observed in a portion of our patients had not been previously described in reported cases, expanding the clinical phenotype of MRD57. These included neurodevelopmental (i.e. difficulties in reading and writing), dysmorphic (i.e. bifid nasal tip and downturned corners of mouth) and skeletal anomalies (i.e. foot polydactyly and short hands). The extended clinical information and comparison with previous literature can be found in (Pavinato et al., 2020).

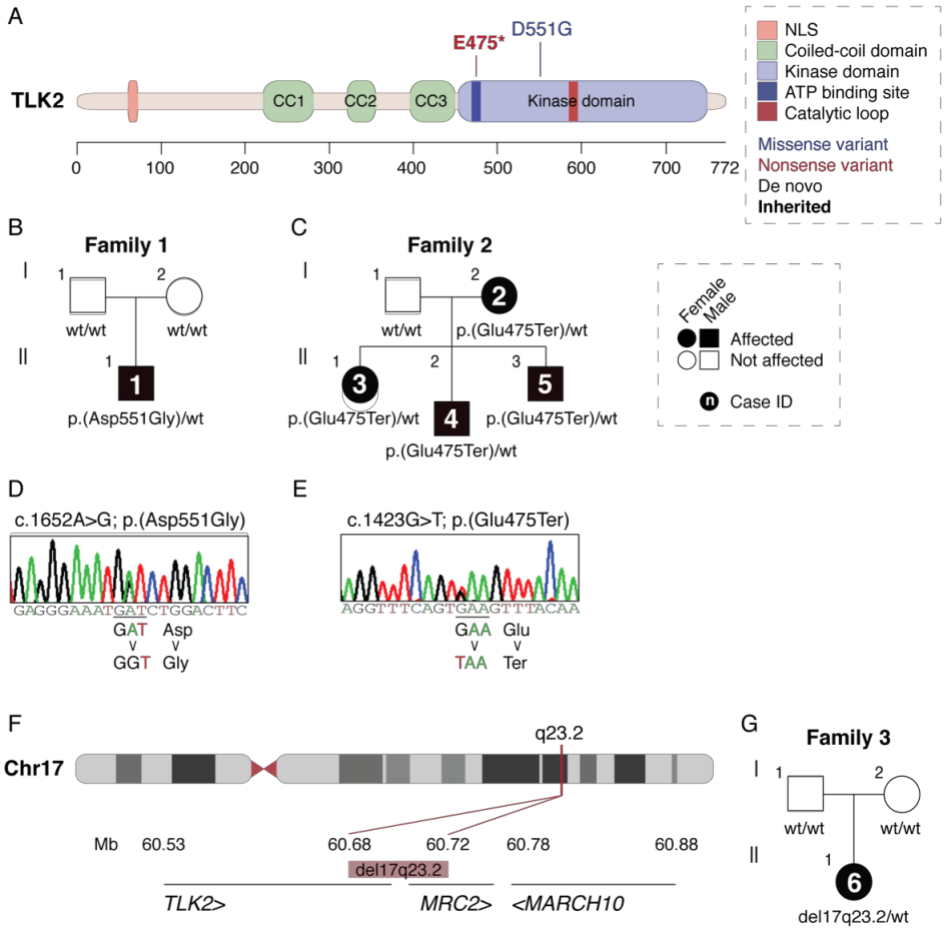


Figure R17. Identification of 6 individuals with novel TLK2 variants.

(**A**) Diagram of TLK2 protein showing the localization of the mutations from families 1 and 2. (**B**, **D**) Case 1 from family 1 carries a heterozygous *de novo* missense variant (c.1652A>G; p.(Asp551Gly)). Pedigree of family 1 (**B**) and Sanger sequencing validation on gDNA of affected cases and unaffected relatives (**D**). (**C**, **E**) Cases 2, 3, 4 and 5 from family 2 share a heterozygous nonsense variant (c.1423G>T; (p.Glu475Ter)) that generates a premature stop codon, inherited from an affected mother. Analysis of maternal grandparents' genotype was not possible, but familial clinical history did not suggest a possible MRD57-like phenotype for them. Pedigree of family 2 (**C**) and Sanger sequencing validation on gDNA of affected cases and unaffected relatives (**E**). (**F**) Diagram of chromosome 17 showing the 17q23.2 deletion (chr17-60683462-60722398) carried by case 6. (**F**-**G**) Case 6 from family 3 carries a heterozygous *de novo* deletion of at least 39 kb, encompassing TLK2 and MRC2 genes. Pedigree of family 3 is shown in (**G**). All variants are referred to the human genome assembly GRCh37 annotation and to NM_001284333.2. wt = wild type at variant position.

TLK2 mutations disrupt proper chromatin compaction

As TLK2 has a known role in controlling chromatin remodeling (Segura-Bayona and Stracker, 2019), we investigated the possibility that the TLK2-D551G mutation causes changes in chromatin compaction. Our collaborators generated patient-derived lymphoblastoid cells (LCLs) from case 1 and performed the single-cell gel electrophoresis (SCGE) assay to address the state of the chromatin in these cells and compared it to age-matched wild-type control LCLs (TLK2 WT in Figure R18).

A relatively short electrophoresis time (20 minutes) was sufficient to unmask slight differences between TLK2 WT and p.(Asp551Gly) cells (Figure R18A). Longer run times (60 minutes) revealed a significantly more relaxed state of the chromatin in LCLs from case 1 compared with WT cells (Figures R18A and R18B). Differences were quantified as 'tail moment' values, which are defined as the product of the tail length and the percentage of DNA in the tail. Furthermore, LCLs from case 1 presented with higher sensitivity to γ-ray irradiation (Figures R18C and R18D) which means they present with increased susceptibility to DNA damage, in line with a more relaxed state of chromatin. Of note, a more relaxed state of chromatin was also observed in fibroblasts derived from case 6, the patient carrying the 17q23.2 deletion encompassing *TLK2* (Pavinato et al., 2020), suggesting a similar chromatin defect in the 17q23.2 deletion and the p.(Asp551Gly) variant. Overall, these findings demonstrated that TLK2 haploinsufficiency disrupts proper chromatin compaction and increases susceptibility to DNA damage.

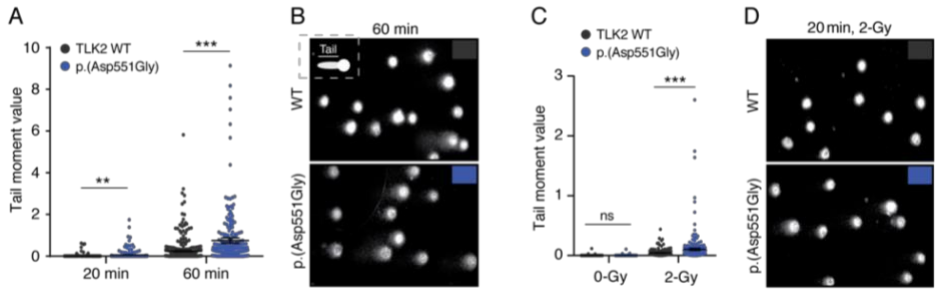


Figure R18. TLK2 mutations disrupt proper chromatin compaction.

(**A**-**B**) Single-cell gel electrophoresis (SCGE) of lymphoblastoid cells (LCLs) carrying the p.(Asp551Gly) mutation and TLK2 wild-type (WT) cells after 20 and 60 minutes of electrophoresis run time. Representative images are shown in (**B**). (**C-D**) SCGE of LCLs carrying the p.(Asp551Gly) mutation and TLK2 WT cells after 20 minutes run time and upon γ-ray irradiation. Representative images are shown in (**D**). DNA migration was quantified as Tail moment values, which are defined as the product between the tail length and the percentage of DNA in the tail. For each condition, at least 100 cells were analyzed. Bars depict mean with SEM of n=3 (**A**-**B**) and n=4 (**C-D**) independent experiments. Statistical significance was determined using an unpaired t-test with Welch's correction (***P<0.001, **P<0.01, ns; not significant).

TLK2 missense mutations alter the activity and subcellular localization of the protein

Most reported missense variants identified in MRD57 patients cluster in the C-terminal PKD of TLK2 (Lelieveld *et al.*, 2016; Reijnders *et al.*, 2018), the core of TLK2 function. Previously, three missense (H493R, H518R, D629N) and one nonsense (R720A) mutation localized in this region (Lelieveld et al., 2016) were functionally tested (Mortuza et al., 2018) (Figure I8). All of these mutations showed a strong decrease in kinase activity, as seen by a reduction in ASF1a phosphorylation *in vitro*, suggesting that these mutations can have a dominant negative effect (Mortuza et al., 2018). However, how these mutations influence TLK2 at the cellular level or what the effect of mutations outside the PKD is remains to be studied. Thus, we decided to investigate how 4 TLK2 mutations identified in MRD57 patients affect TLK2 interactions, localization and activity *in vitro*.

The mutations we chose are located in different regions of the TLK2 protein. Two of them, p.(Asp551Gly) (hereafter D551G) and p.(Ser617Leu) (hereafter S617L), are *de novo* missense variants in heterozygosis, and are located within the PKD (Figure R19A). The D551G mutation was identified in our survey (Figures R17A, R17B and R17D), while the S617L mutation was found in a patient with ASD described in previous WES works (O'Roak et al., 2011, Iossifov et al., 2014, Satterstrom et al., 2020). This mutation was also reported in the ASC exome analysis browser (https://asc.broadinstitute.org/), a freely accessible portal containing de novo variants identified in affected patients enrolled in ASC WES

projects. The S617 residue was of particular interest because it is located one residue after the Asp-Phe-Gly (DFG) motif (Treiber and Shah, 2013, Vijayan et al., 2015), that together with the activation and catalytic loop, constitute the kinase core of TLK2 (Mortuza et al., 2018). We previously identified S617 as a TLK2 auto-phosphorylation site and shown that alterations of this residue are able to enhance or impair TLK2 kinase activity, depending on the amino acid change (Mortuza et al., 2018). Both mutations were predicted as likely pathogenic by several *in silico* predictors (Figure R19D). The other two mutations, p.(Lys55Glu) (hereafter K55E) and p.(Gly297Asp) (hereafter G297D), have been previously published and are located outside the PKD (Reijnders et al., 2018, Topf et al., 2020) (Figure R19A). K55E is a hereditary missense variant and the only homozygous mutation in TLK2 described to date (Topf et al., 2020). The mutation has been predicted to be damaging by *in silico* tools. The parents were found to be heterozygous carriers but clinically unaffected, suggesting that TLK2 variants can cause a recessive disease, with a more severe phenotype than previously described in MRD57 patients (Topf et al., 2020). In the case of G297D, it is a *de novo* missense variant in heterozygosity, and the mutation is located outside any known functional domain of TLK2, between the CC1 and CC2 coil-coiled domains. The mutation affects a highly conserved amino acid and was also predicted to be pathogenic by several *in silico* tools (Reijnders et al., 2018).

We first examined the potential structural impact of the D551G and S617L mutations by modelling them in the TLK2 PKD crystal structure (Mortuza et al., 2018). The D551G mutation was predicted to weaken hydrogen bonds with the subsequent helix (Figure R19B), while S617L introduced a hydrophobic residue in place of the auto-phosphorylation site (Figure R19C). We could not model the K55E or G297D mutations, as we do not yet have the crystal structure of the N-terminus. However, K55E exchanged a positive for a negatively charged side chain, while G297D changes a glycine to a larger, charged amino acid.

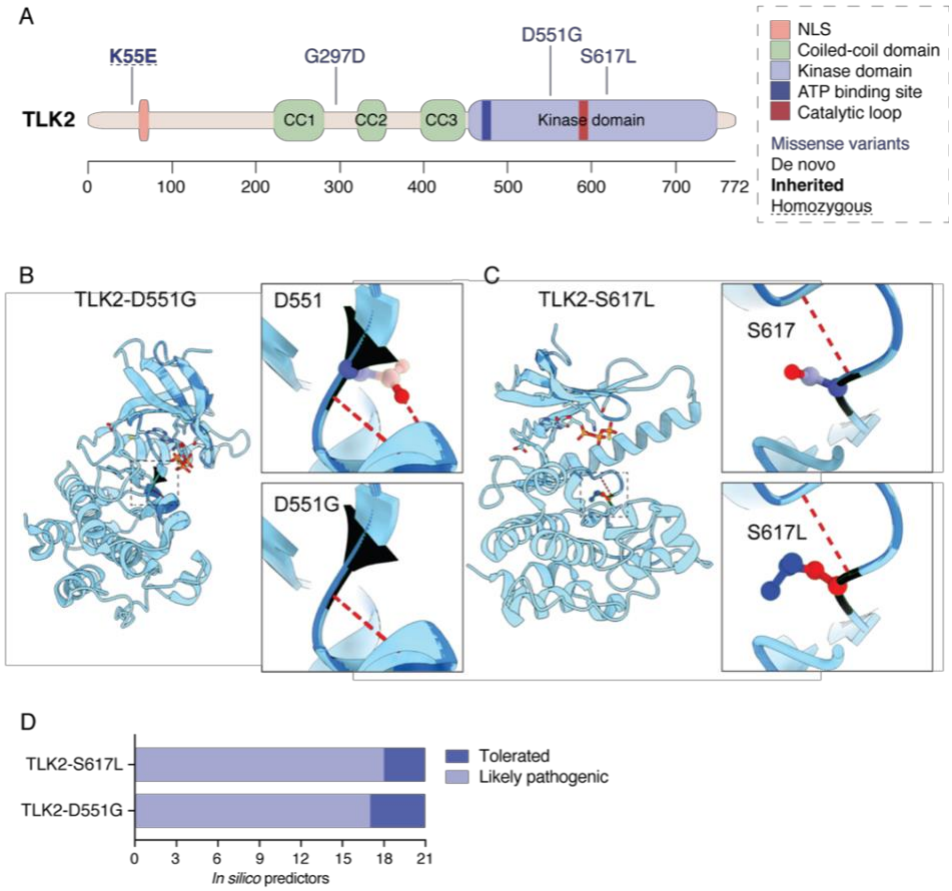


Figure R19. Modelling TLK2 missense mutations.

(**A**) Diagram of TLK2 protein showing the localization of the 4 mutations. D551G and S617L are located within the protein kinase domain (PKD) while K55E and G297D are located outside the PKD. (**B**-**C**) Modelling of the D551G (**B**) and S617L (**C**) missense mutations on the crystal structure of the TLK2 PKD (Mortuza et al., 2018). Hydrogen bonds are shown in red dashed lines. (**D**) *In silico* prediction of the functional impact of D551G and S617L.

In order to perform *in vitro* experiments, we generated the corresponding patient mutations via site-directed mutagenesis in the Strep-FLAG tagged vector pcDNA3.1 N-SF-TAP-TLK2-WT (Mortuza et al., 2018). We confirmed the mutations by Sanger sequencing and checked the overexpression by WB. To determine if these mutations impaired kinase activity, we performed *in vitro* kinase assays (as in Figure R9). To do so, we first transiently expressed in AD-293 cells Strep-FLAG tagged empty vector (SF-TAP), TLK2-WT, TLK2 kinase dead

(TLK2-KD, D592V mutation) and TLK2 missense variants (TLK2- D551G, S617L, K55E, and G297D). We then immunoprecipitated the proteins with the N-terminal strep tag and incubated with purified substrate (ASF1A) and radioactive ^{32}P for kinase assays (Figure R20A).

The mutations in the PKD, D551G and S617L, showed higher expression compared to TLK2-WT, which we have previously observed in TLK2 mutants that have reduced kinase activity (Mortuza et al., 2018). Indeed, D551G and S617L led to a significant reduction in substrate modification (Figure R20A). Quantification of multiple experiments showed that TLK2-S617L impaired kinase activity to levels comparable to the TLK2-KD control (Figure R20B). TLK2-D551G results were more variable and showed a milder substrate modification reduction and slightly higher autophosphorylation levels than TLK2-WT in some experiments (Figure R20B). On the other hand, the mutations located outside the PKD, K55E and G297D, surprisingly showed a significant increase of both substrate and auto phosphorylation levels (Figure R20A). Both mutations appear to be hyperactive, with G297D autophosphorylation levels being 5-15 times those of TLK2-WT (Figure R20C).

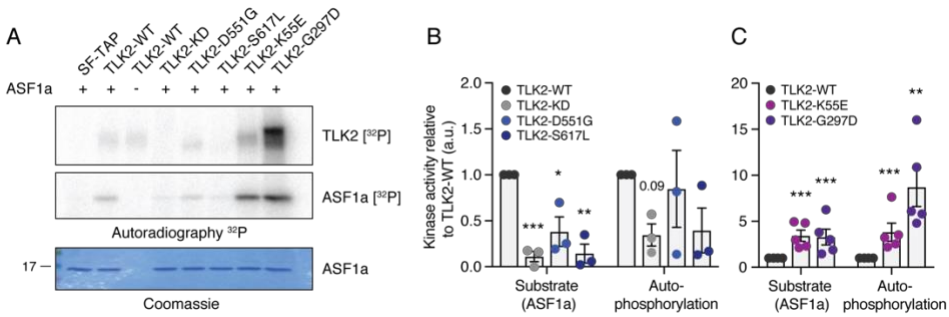


Figure R20. TLK2 missense mutations alter the activity of the protein.

(**A**) Representative *in vitro* kinase assays with Strep-purified TLK2-WT (wild-type), TLK2-KD (kinase dead; D592V) and the indicated missense variants. Autophosphorylation (TLK2) and substrate (ASF1A) phosphorylation are shown. Coomassie is shown as loading control for ASF1A. (**B-C**) Quantification of n=3 (**B**) and n=5 (**C**) independent kinase assays. Individual results (dots) are shown for each assay on purified ASF1A substrate or affinity purified TLK2 autophosphorylation relative to corresponding TLK2-WT. Bars depict mean with SEM. ANOVA followed by Dunnett's test was applied when comparing multiple experimental groups against a control (TLK2-WT). Statistical significance was considered when adjusted p-values were ***P<0.001,

**P<0.01, *P<0.05. All tests were performed after log-transformation of the data. However, for clarity of representation, data are shown in the original scale. Statistical analysis was carried out with the multcomp (Hothorn et al., 2008) R package.

TLK2 is a nuclear kinase that exhibits a diffuse nuclear localization. In previous work, we observed that deletion of any of the N-terminal coiled-coil domains of TLK2 led to perinuclear localization (Mortuza *et* al., 2018). To determine if the ASD missense mutations altered TLK2 localization, we transfected AD-293 cells with Strep-FLAG tagged TLK2-WT and ASD mutants and performed immunofluorescence (IF). We used Lamin A as an inner nuclear membrane marker, stained nuclear DNA with DAPI and TLK2 with FLAG antibody. TLK2-WT showed diffuse nuclear localization in transfected cells as expected (Figure R21A and R21B). In contrast, TLK2-D551G and TLK2-S617L showed perinuclear accumulation to different extents (Figure R21A and R21B). This was particularly prominent for the TLK2-S617L mutant, where 75% of transfected cells showed perinuclear localization of TLK2 (Figure R21B). We suspected that TLK2-K55E would disrupt the predicted NLS of TLK2. However, TLK2-K55E showed a normal diffuse nuclear localization, while TLK2-G297D had a slight perinuclear localization, but much milder than TLK2-D551G and TLK2-S617L (Figure R21B).

Based on previous data, the main pathogenic mechanism of *TLK2* mutations appeared to be a reduction in overall activity (Mortuza et al., 2018, Reijnders et al., 2018). Our results indicate that not all the missense mutations in MDR57 patients are loss-of-function and suggest that the mechanism by which TLK2 mutations cause NDD is not exclusively due to a reduction in activity, but rather an imbalance in the levels of TLK2 activity. Our results also suggest that the localization of TLK2 may be linked to the proper regulation of its kinase activity and its mislocalization may also play a role in the molecular defects underlying the pathology of MRD57.

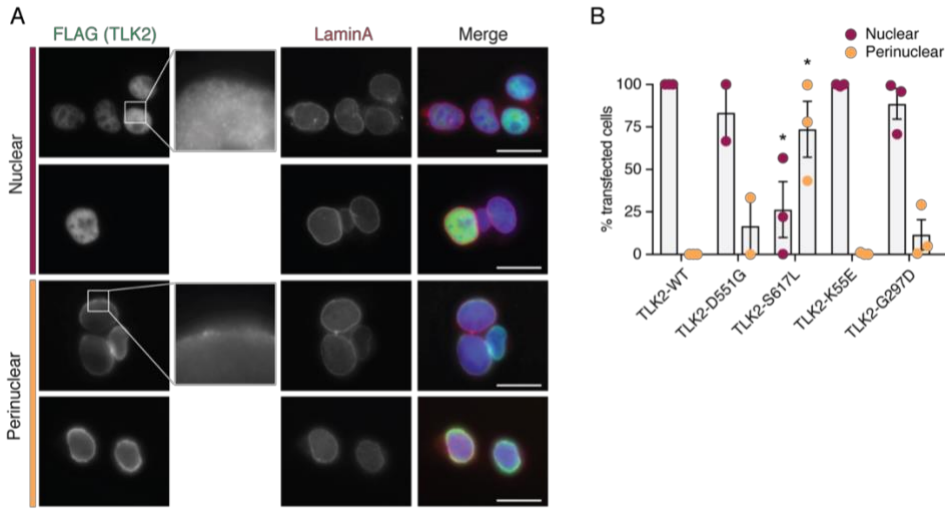


Figure R21. Missense mutations alter the subcellular localization of TLK2.

(**A**) Representative IF microscopy of overexpressed TLK2 in AD-293 cells, indicating the two main localization patterns observed. The nuclear localization image corresponds to TLK2-WT, while the perinuclear localization to TLK2-S617L. Scale bar=20 μM. (**B**) Quantification of TLK2 localization patterns for TLK2-WT and the indicated missense variants. Ten random fields were scored in n=2 (TLK2-D551G) or n=3 (TLK2-WT, S617L, K55E, and G297D) biological replicates. Bars depict mean with SEM. Statistical significance was determined using an unpaired t-test with Welch's correction (*P<0.05).

The TLK2 interactome contains proteins involved in neurodevelopmental disorders

Next, we wanted to asses if TLK2 interacted with proteins known to have roles in NDDs and if the missense mutations altered the proximal interactome of TLK2. To do so, we used a proximity-dependent biotin identification assay (BioID) (Roux et al., 2012) coupled to mass spectrometry (MS) (BioID-MS). We tagged TLK2-WT with the promiscuous *E.coli* biotin protein ligase (BirA*) and expressed the fusion protein in AD-293 cells. The expression of BirA* in the presence of biotin leads to the biotinylation of proteins in the proximity of the fusion protein (proximal interactors). After stringent cell lysis and protein denaturation, the biotinylated proteins can be affinity purified and identified by mass spectrometry (LC-MS/MS) (Figure R22A). As compared to other methods, BioID has the advantage that it does not rely on maintaining delicate complexes

or high affinity interactions that often do not take place between kinases and their substrates, thus allowing us to identify not only strong interactors, but also proteins that are in close proximity to our bait.

The missense mutations were generated via site-directed mutagenesis in the BirA*-N-term-TLK2-WT plasmid. BirA*-tagged TLK2-WT, TLK2-D551G, TLK2-S617L, TLK2-K55E, TLK2-G297D and empty vector were expressed in AD-293 cells by transient transfection, biotin was added, and biotinylated proteins were pulled down by Streptavidin-IP from cell lysates and sent for MS analysis (Figure R22B). Confirmation of overexpression and successful biotinylation was shown by WB (Figure R22C).

TLK2-WT network clustering identified the known TLK2 substrates ASF1A, ASF1B and TLK1, as well as DYNLL1/2 (LC8) that we have previously validated as an interactor but that does not appear to be a TLK2 substrate (Figure R22D) (Segura-Bayona et al., 2017). The TLK2-WT proximal interactome grouped into five functional clusters consistent with the known functions of TLK2, including chromatin binding or remodeling, DNA repair and histone chaperones, RNA processing and splicing, and transcriptional regulation (Figure R22D).

We then cross-referenced the TLK2-WT proximal interactome with a recent large-scale analysis of iPOND-MS data that identified TLK1 and TLK2 as high confidence interactors with nascent DNA at active replication forks (Wessel et al., 2019). Several of the TLK2-WT hits, including ASF1B, CHD8, RAD50, BRD4, SCML2, and NACC1, were also found at active forks. In order to investigate if TLK2 was in close proximity to proteins known to be involved in NDDs, we compared our TLK2-WT proximal interactome to the Simon's Foundation Autism Research Initiative (SFARI) and DECIPHER databases, which respectively contain information on autism-susceptibility genes and genomic variants that cause rare diseases (Firth et al., 2009, Banerjee-Basu and Packer, 2010, Abrahams et al., 2013). We identified 8 high confidence hits: CHD7, CHD8, RAD50, NACC1, CCNK, JMJD1C, MSANTD2, and YEATS2, that were in one of these 2 databases. These data suggest a potential functional link between TLK2 and a number of proteins known to be involved in NDDs with overlapping pathologies. Further details about these BioID-MS results, and

SFARI and OMIM classifications of the TLK2 proximal interactors can be found in the Supplementary table 3 and in (Pavinato et al., 2020).

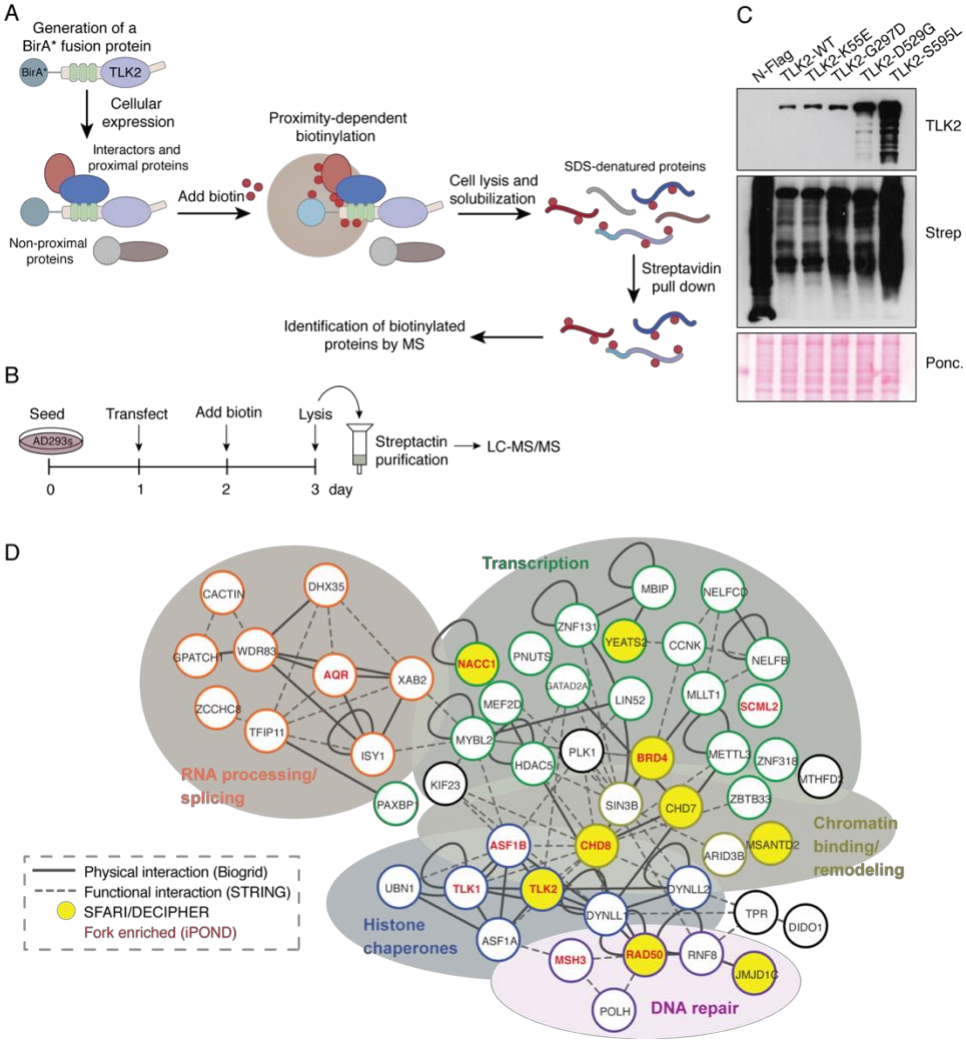


Figure R22. The TLK2 interactome contains proteins involved in neurodevelopmental disorders.

(**A**) Principle of BioID-MS. The expression of a promiscuous *E.coli* biotin protein ligase (BirA*) fused to a target protein in live cells leads to a proximity-dependent biotinylation of proteins that are close to the fusion protein. Biotinylated proteins can be affinity purified and identified by MS. Adapted from (Roux et al., 2012). (**B**) Schematic of the BioID-MS experiment performed. TLK2 plasmids were transfected in AD-293 cells, biotin was added 24 hours later and cell lysed 24 hours after that. Biotin-labeled proteins were then Strep-purified and send to MS for analysis. (**C**) Western blot analysis of AD-

293 lysates expressing BioID constructs: N-FLAG-BirA alone or fused to the indicated TLK2 allele. Detection with anti-TLK2 and streptavidin-horseradish peroxidase are shown. Ponceau stain is shown as a loading control. (**D**) Network clustering of all prey hits with a SAINT score of ≥0.7 in TLK2-WT samples. Physical interactions reported in Biogrid (solid lines) and functional interactions reported in STRING (dashed lines) are indicated (Oughtred et al., 2019, Szklarczyk et al., 2019). Functionally enriched clusters are indicated by color coding. Bait, TLK2 substrates, proteins found in the Simon's Foundation Autism Research Initiative (SFARI)/DECIPHER gene databases (yellow fill) (Firth et al., 2009, Banerjee-Basu and Packer, 2010, Abrahams et al., 2013) or proteins enriched on nascent DNA/replication forks (red font) are indicated (Wessel et al., 2019).

Missense mutations alter the proximal interactome of TLK2

In parallel to TLK2-WT, we performed BioID analysis of BirA*-TLK2-D551G, BirA*-TLK2-S617L, BirA*-TLK2-K55E and BirA*-TLK2-G297D (Figure R22C). Also in this case, both BirA*-TLK2-D551G and BirA*-TLK2-S617L accumulated to higher levels than BirA*-TLK2-WT, consistent with our previous observations with FLAG tagged variants. That was not the case with BirA*-TLK2-K55E and BirA*-TLK2-G297D, which accumulated to similar levels of TLK2-WT (Figure R22C).

All mutations caused numerous alterations in the proximal interactome compared to TLK2-WT, although to different extents (Figure R23). We did not lose in any case the known TLK2 interactors, TLK1, ASF1A/B or LC8. However, we could see a higher level of labeling of TLK1 and the ASF1A/B proteins by TLK2-D551G and TLK2-S617L (Pavinato et al., 2020). This was not unexpected, as we previously observed higher binding in TLK2 mutants with reduced kinase activity.

In the case of the inactivating mutations, TLK2-D551G and TLK2-S617L, we observed a reduction in several replication fork and ASD-related proteins, including RAD50 and YEATS2 with both mutants, and JMJD1C, BRD4, CCNK, NACC1, MSANTD2 and CHD8 with TLK2-S617L only (Figure R23B). In addition, some ASD-related proteins found in the SFARI database, including NFIA, NFIX and PAPOLG, were significantly enriched with both mutants, but not in TLK2-WT (Figure R23A and R23B). At the same time, TLK2-S617L also

showed enrichment of FOXC1, ALX4, EBF3 and HOXD13 proteins (Figure R23B). Mutations in the genes encoding for these proteins have been annotated in the DECIPHER database in patients that present with similar phenotypes to those with MRD57, including ID and microcephaly.

The hyperactivating mutations TLK2-K55E and TLK2-G297D did not produce strong alterations in the TLK2 interactome compared to the inactivating mutations. TLK2-G297D showed enriched interactions with many ASD-susceptibility genes from SFARI including CHD7, CHD8 and NSD1, genes found in patients with ID in the DECIPHER database such as DHX34, OGT and ZNF148, and many fork enriched proteins (Figure R23D). Of interest, YWHAZ, a 14-3-3 protein, was strongly enriched in TLK2-G297D compared to any other TLK2 mutants tested. In the case of K55E, no major changes in the interactome were observed by BioID-MS (Figure R23C).

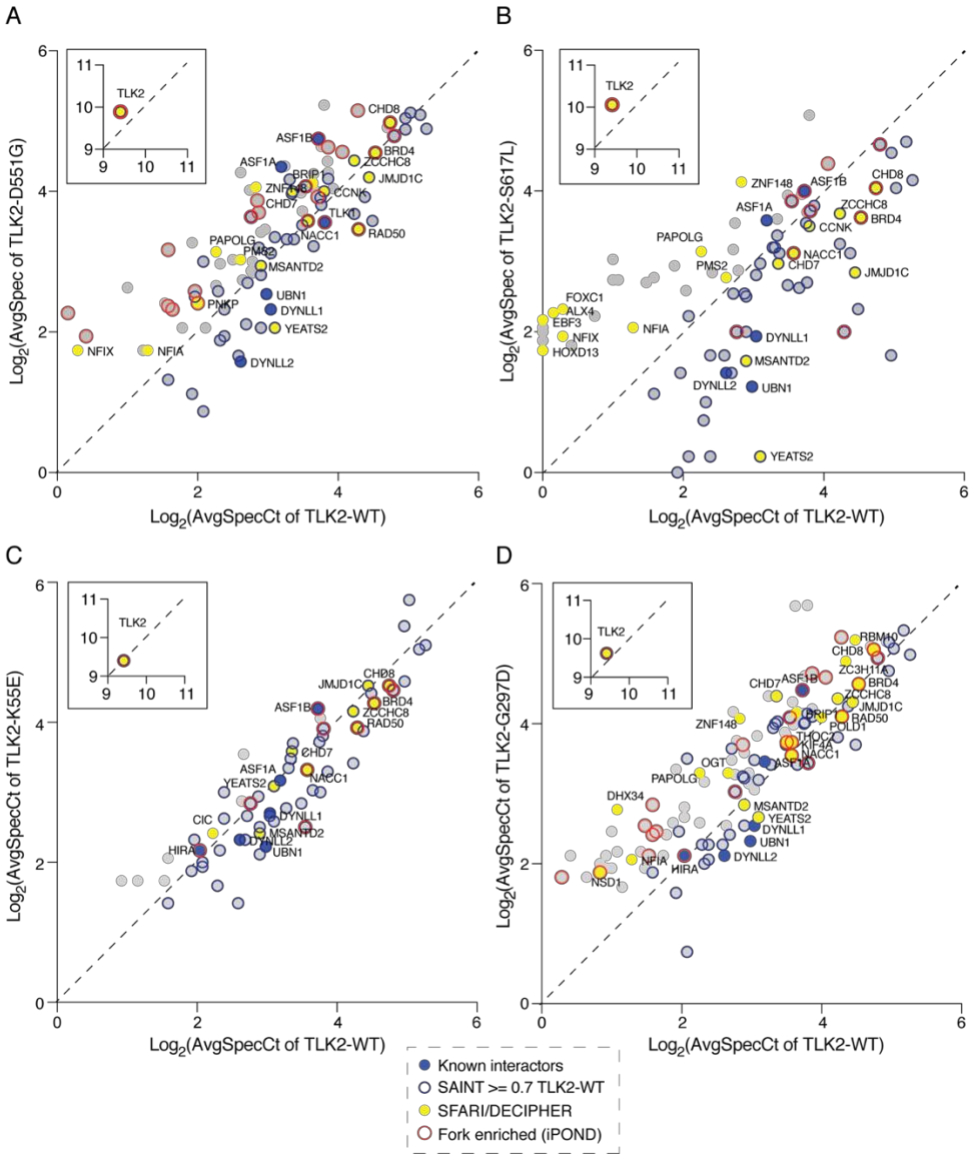


Figure R23. Missense mutations alter the proximal interactome of TLK2 #1.

(**A**-**D**) Scatter plots of average spectral counts (Log_2 transformed) of bait and prey proteins identified with the TLK2-D551G (**A**), TLK2-S617L (**B**), TLK2-K55E (**C**), and TLK2-G297D (**D**) alleles compared to TLK2-WT. Previously identified TLK2 interactors (blue dots), proteins with a SAINT score of ≥0.7 in TLK2-WT (blue circles), proteins enriched on replication forks (red circles) or found in the SFARI/DECIPHER databases (yellow dots) are indicated.

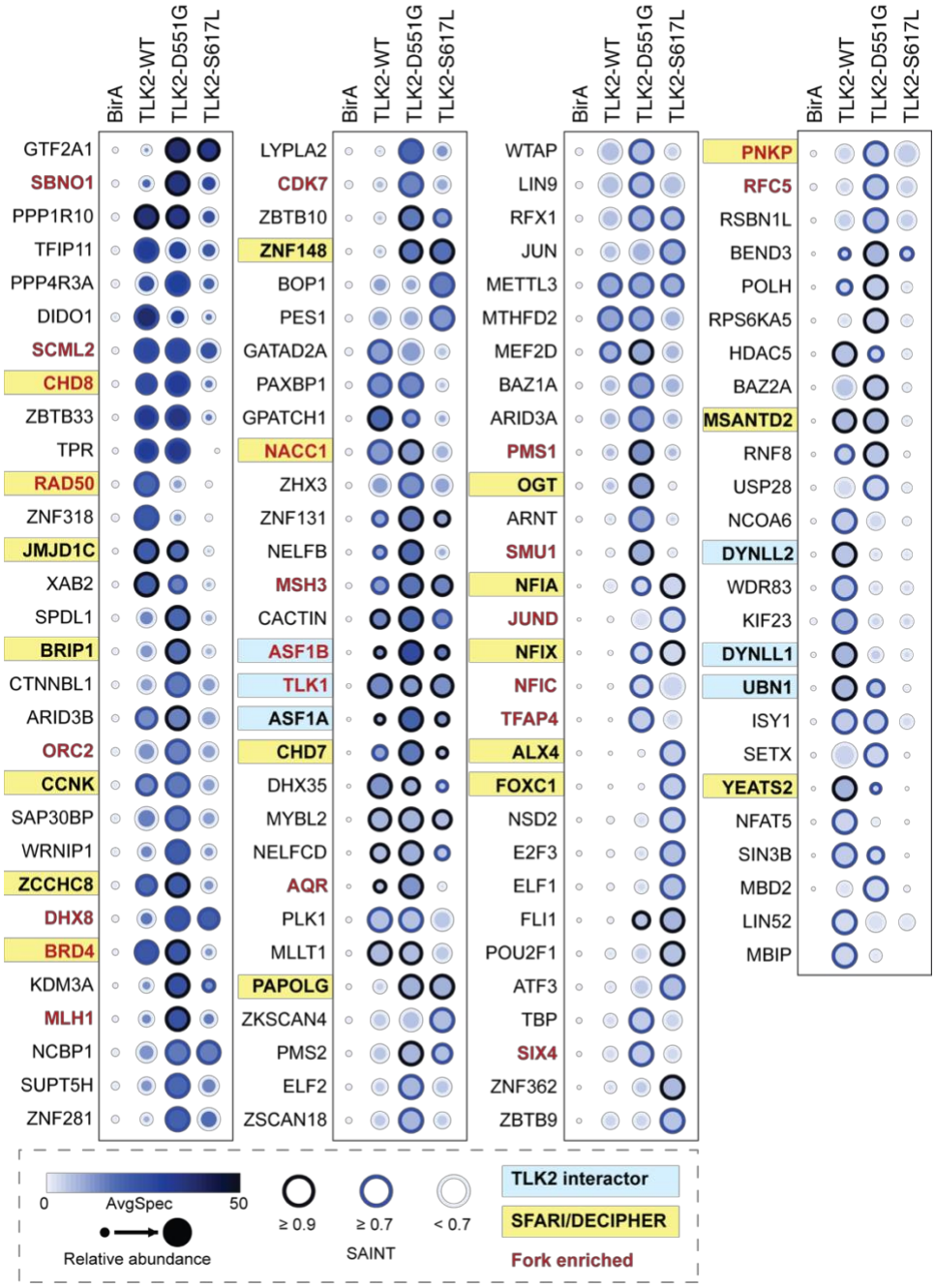


Figure R24. Missense mutations alter the proximal interactome of TLK2 #2.

Dot plot of prey proteins with a SAINT score of ≥0.7 with TLK2-WT, TLK2-D551G or TLK2-S617L as baits generated using ProHits-viz (Knight et al., 2017). Average spectral counts (SC), relative abundance and SAINT score ranges are indicated, as well

as proteins enriched on replication forks (Wessel et al., 2019) or found in the SFARI/DECIPHER databases (see legend).

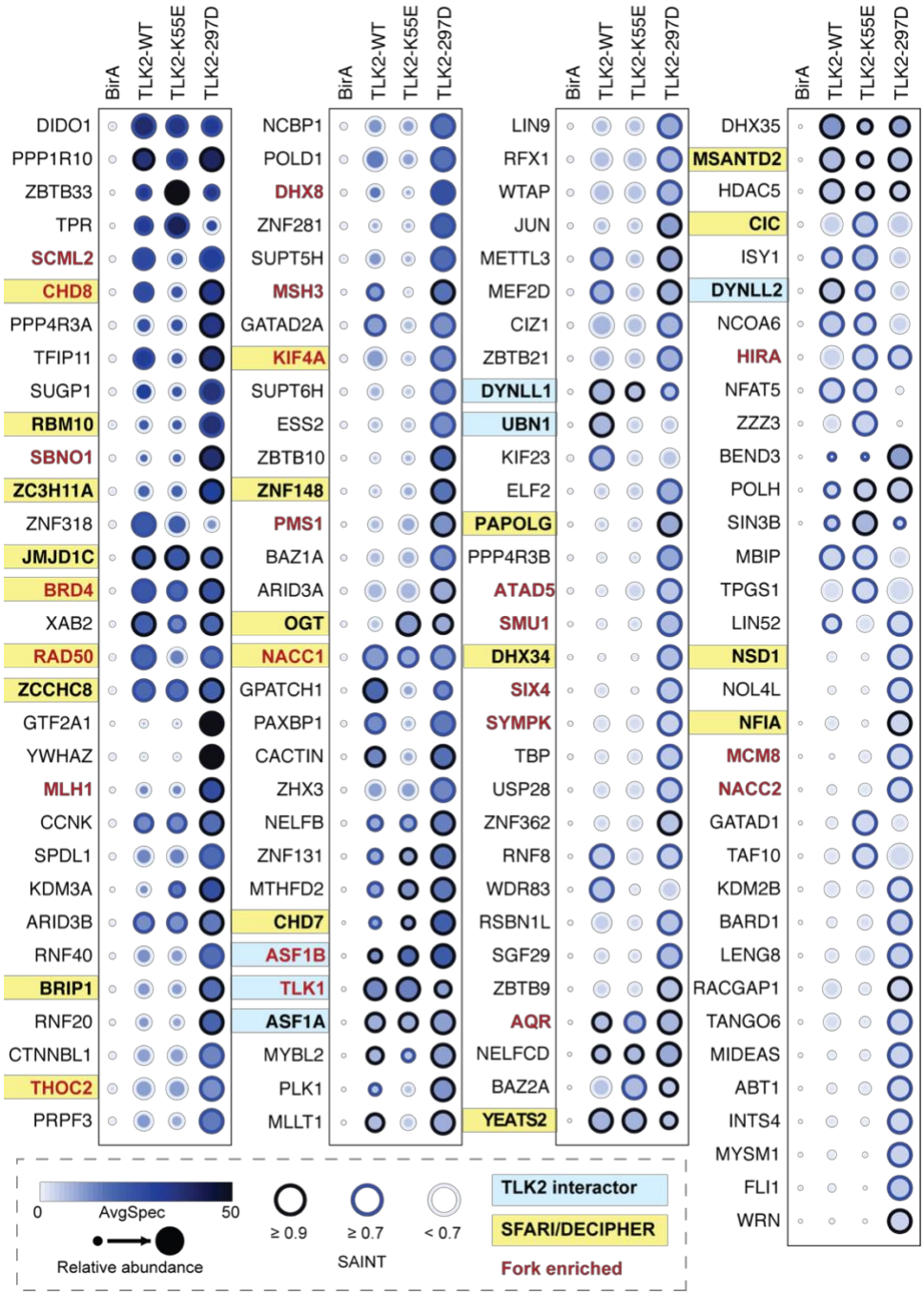


Figure R25. Missense mutations alter the proximal interactome of TLK2 #3.

Dot plot of prey proteins with a SAINT score of ≥0.7 with TLK2-WT, TLK2-K55E or TLK2-G297D as baits generated using ProHits-viz (Knight et al., 2017). Average spectral counts (SC), relative abundance and SAINT score ranges are indicated, as well as proteins enriched on replication forks (Wessel et al., 2019) or found in the SFARI/DECIPHER databases (see legend).

We performed BioID in a small-scale format followed by WB to validate a few of these proximal interactions and their relative differences. We first validated the known interactors (TLK1, ASF1, and LC8), and then we tested some of the top hits from our BioID-MS (CHD7, CHD8, RAD50, NBS1, MRE11, BRD4, and MTHFD2) (Figures R22D and R26). Expression of the control N-FLAG-BirA (BirA* empty) supplemented by biotin led to no detection of any of the aforementioned proteins, suggesting that the interaction is specific to TLK2. All the proteins were clearly validated with all TLK2 alleles, although in the case of TLK2-K55E signal was reduced, most likely because this mutant was expressed at lower levels. Interestingly, in the case of TLK2-S617L, LC8, CHD7, and CHD8 signal was reduced, as well as MRN complex proteins RAD50 and NBS1, consistent with the MS data (Figure R26).

Another proximal interactor that caught our attention is replication timing regulatory factor 1 (RIF1). RIF1 has come out in in all our MS experiments, including IP-MS and BioID (Segura-Bayona et al., 2017), although always below our stringent statistical threshold due to its appearance in a subset of controls. RIF1 has critical roles in both replication timing control and DSB repair pathway choice (Ira and Nussenzweig, 2014), and several MS studies have previously identified RIF1 as a potential interactor of TLK1/2 and ASF1a (Sukackaite et al., 2017, Yang et al., 2018). We set out to validate this possibility by transfecting AD293 cells with empty, TLK2, ASF1a, and ASF1b SF-TAP constructs and performing WB analysis after immunoprecipitation. Expression of the control SF-TAP vector led to no detection of RIF1 (Figures R27A and R27B), validating that RIF1 specifically interacts with TLK2 (Figure R27A) and both ASF1a and ASF1b (Figure R27B). However, whether TLK2 interacts directly with RIF1 or does so via ASF1a/b remains to be determined.

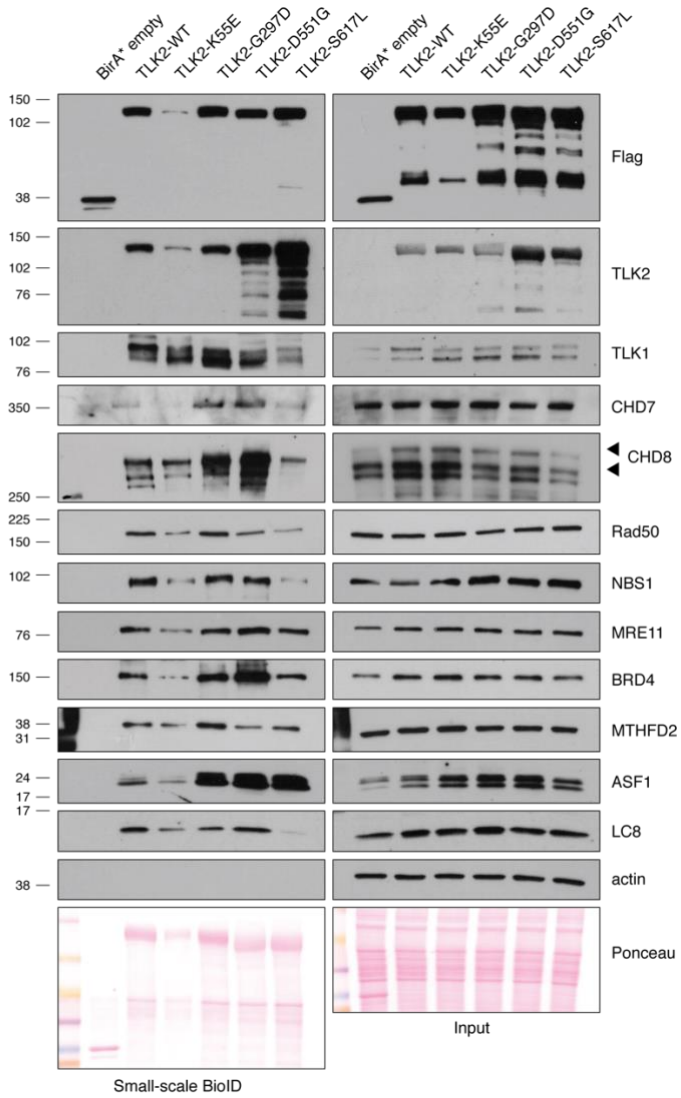


Figure R26. Validation of proximal interactors by BioID-WB.

Western blot analyses of the indicated proteins from Strep-affinity purification (AP) lysates from AD-293 cells transfected with the indicated BioID construct and supplemented with biotin. Input levels are shown and a representative ponceau staining is provided as a loading control. Actin was used as a loading control for the input and as a negative control for the small-scale BioID. Representative of n=2 independent experiments.

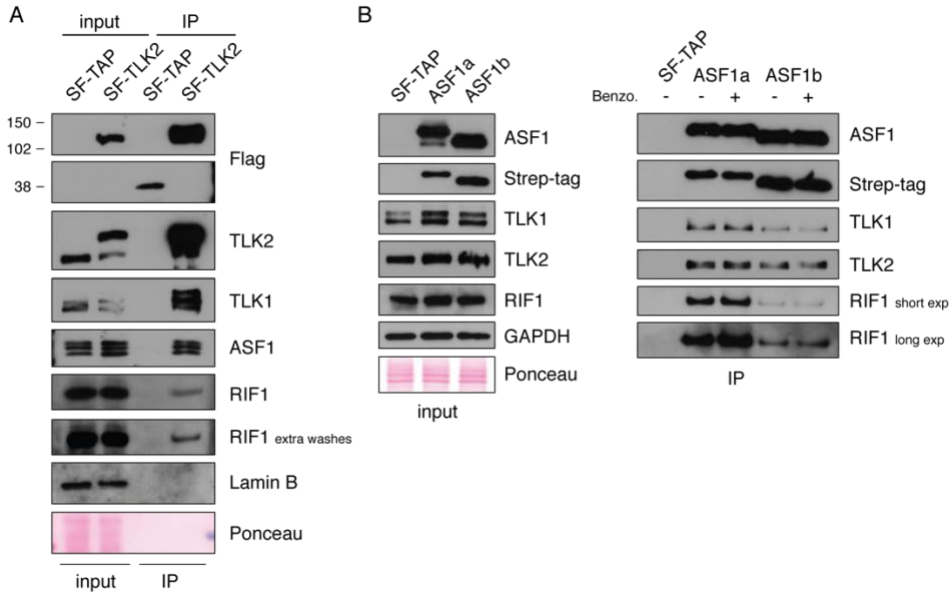


Figure R27. Validation of proximal interactors by IP-WB.

(**A**-**B**) Western blot analyses of the indicated proteins from Strep-affinity purification (AP) lysates from AD-293 cells transfected with the indicated SF-TAP construct. Input levels are shown and a representative ponceau staining is provided as a loading control. Lamin B and GAPDH were used as additional loading controls. In (B), benzonase was used to degrade all DNA and RNA. Representative of n=3 (**A**) and n=2 (**B**) independent experiments. Benzo, benzonase.

Discussion

In this thesis, we examined the *in vivo* roles of TLK2 in the context of disease and immune suppression ermerged as a common feature in TLK2's role in both cancer and NDDs. In cancer, high levels of TLK2 contribute to the suppression of the immune system, potentially enhancing cancer aggressiveness. This could be exploited therapeutically, as TLKs are druggable targets expressed in all cancers and their inhibition could potentially boost an immune response and sensitize tumors to ICB. Similarly, mice lacking TLK2 showed increased transcription of inflammatory genes and signatures of glial components in the brain. This has been previously reported in the brains of patients with several NDDs, suggesting that inflammation could be a major contributor to the pathogenesis of MRD57.

TLKs as a potential cancer target to augment immune checkpoint blockade efficacy

In the results presented in Chapter 1, we showed how depleting TLK2 in already formed breast tumors stalled primary tumor growth and reduced the incidence of lung metastasis in an MMTV-PyMT genetic model (Figures R3 and R4). In previous *in vitro* work, we observed that the depletion of both TLK1 and TLK2 in transformed mouse embryonic fibroblasts (MEFs) or human cancer cell lines, there was a reduction in colony formation capacity, and an increase in DSBs, micronuclei, and other chromosomal aberrations (Segura-Bayona et al., 2017, Lee et al., 2018). Similarly, when we derived cell lines by spontaneous immortalization from MMTV-PyMT *Tlk2*$^{F/F}$ UbCre^{ERT2} tumors and induced TLK2 depletion by *in vitro* 4-OHT treatment, cells showed a decrease in colony formation capacity and BrdU incorporation accompanied by an increase in SSBs, DSBs and micronuclei (Segura-Bayona, 2018). We therefore initially hypothesized that TLK2 loss would impair breast cancer progression due to this cell autonomous effect. The tumor cells would proliferate less due to stalled forks and increased levels of DNA damage.

However, this does not appear to be the case when the tumor cells are in a physiological context, surrounded by, and interacting with, the many cell types that form the TME. 15 days post 4-OHT treatment, we observed a very mild decrease in proliferation (Ki67 staining) within the tumors (Figures R3B and

R3C) and an absence in gene expression changes in genes related to proliferation, cell cycle, apoptosis, or senescence. We also could not readily observe an increase in γH2AX in the *Tlk2*$^{-/-}$ tumors with the exception of γH2AX^{+} signal that coincided with apoptotic areas (Cleaved Caspase 3^{+}) in large tumors, regardless of TLK2 status. This is consistent with the fact that depletion of both TLK1 and TLK2 in cancer cells *in vitro* is required for a strong DDR in most cell types tested.

Our analysis of genomic data from nearly 15,000 patients in TCGA datasets revealed that high levels of *TLK1* and *TLK2* correlated with increased CIN, low *STING* levels and low immune infiltration in many cohorts (Figures R1A and R1B, Supplementary table 1). In a small subset of those TCGA datasets, *TLK2* levels anti-correlated with signatures of CD8^{+} cytotoxic T cells, activated NK cells, and M1 macrophages. With a few exceptions, these immune subtypes are considered to have positive prognostic value, as their proinflammatory properties drive an antitumoral response (Extensively reviewed in Bruni et al., 2020). The correlation with *STING* levels and immune infiltration was stronger in the case of *TLK2*, which is also more frequently amplified in cancer (Lee et al., 2018). In some cases, such as cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) and Liver Hepatocellular Carcinoma (LIHC), high *TLK2* expression correlated with both reduced levels of STING and stromal infiltration (Figures R1A and R1B) (Segura-Bayona et al., 2020), as well as with decreased disease-free survival (Figure I6) (Lee et al., 2018). This is in line with immune suppressed or *cold tumors* generally having worse prognosis (Pages et al., 2010, Barnes and Amir, 2017, Ye et al., 2019, Bruni et al., 2020).

Although the characterization of the intra-tumoral immune infiltrates presented in Chapter 1 is preliminary and lacks statistical power, we observed shifts in key immune populations, such as an increased recruitment of monocytes and immature TAMs (Figure R6G and R6J) and an increase in NK cells and activated T cells expressing higher PD-1 and CD69 (Figure R7C and R7H-P) at an early time-point following TLK2 deletion. In future work, we would like to perform a more in-depth analysis of the immune subsets, and also look at memory and regulatory T cells, characterize the polarization of the TAMs, analyze the cytokines and chemokines in the milieu and what cell types produce them, and assess the expression of PD-L1 in the tumor cells. Such characterization could be performed by mass cyometry (CyTOF) of the tumors or by single-cell RNA-seq of the CD45^{+} cells.

One important limitation of our system is that we used a UbCre, and therefore all the cell types of the mice become *Tlk2*$^{-/-}$ upon 4-OHT treatment. This makes it impossible for us to discriminate between the effect of knocking out TLK2 in the tumor cells versus other cell types, such as macrophages and T cells, which are known to play a role in mammary tumorigenesis and metastasis in this particular model (Lin et al., 2001, DeNardo et al., 2009). A possibility is that when the tumor cells become *Tlk2*$^{-/-}$ they present with spurious transcription of repetitive elements and an antiviral and immune response, as we have observed in cells *in vitro* (Segura-Bayona et al., 2020). This could culminate in an increase in IFN production or other pro-inflammatory cytokines that attract and recruit the immune system to the tumors. The IFNγ response usually happens very quickly in tumors which would explain the timing of our phenotype. Another possibility is that knocking out TLK2 has an effect on the activity of immune cells. Based on our initial characterization of TLK2 deficient mice, we have no indication that there are major deficits in immune system function or development, but it remains a possibility that there are more subtle issues relevant to the phenotype (Segura-Bayona et al., 2017). There are several experiments that we plan to do in the future to address this very important point (discussed below).

The results described above suggest that *Tlk2*$^{-/-}$ tumors could respond to ICB treatment. ICB has revolutionized cancer treatment for some patients and is given as a first line of treatment in patients with metastatic melanoma, as well as hepatocellular carcinoma. However, most patients do not show long-lasting remission, while some remain completely refractory to the treatment (Sharpe and Pauken, 2018). Efforts to define biomarkers that predict which patients will benefit from ICB have elucidated that the presence of CD8$^+$ T cells, high levels of PD-L1 expression in the tumor, and an ongoing inflammatory milieu, as well as high mutational burden of the tumor cells, usually correlate with a positive clinical outcome (Brahmer et al., 2012, Topalian et al., 2012). Therefore, finding therapies that increase these biomarkers is important to boost ICB efficiency. Currently, many clinical trials are underway combining αPD-1 antibodies with other therapies, which include blocking other inhibitory receptors and immunoregulatory cytokines, and giving agonists for stimulatory molecules (Sharma and Allison, 2015).

Our hypothesis that TLK inhibition might boost ICB efficacy is supported by a recent study that showed that loss of the downstream target of TLKs *Asf1a* sensitizes *Kras*-mutant lung adenocarcinoma (LUAD) to anti–PD-1 treatment by polarizing macrophages towards M1 and enhancing T cell activation (Li et al., 2020). While TLK1 and TLK2 were not in the CRISPR library used in the study, we imagine their individual or combined loss may yield a similar effect, a hypothesis we will test in the near future using similar systems.

Both ASF1a and ASF1b are upregulated in other cancer types and correlate with poor prognosis (Corpet et al., 2011, Wu et al., 2019). Efforts have been made to generate ASF1 inhibitors (Miknis et al., 2015, Richet et al., 2015, Seol et al., 2015), and a peptide that efficiently dissociates the interaction between ASF1s and the histone H3-H4 heterodimer shows some effect in reducing tumor growth *in vivo* (Bakail et al., 2019). Due to the central role that ASF1s play in histone supply, and because peptide-based drugs are harder to efficiently deliver, we believe that TLKs are potentially a better target but specific small molecules remain to be identified.

Although several potential TLK inhibitors have been described in the literature (Ronald et al., 2013, Kim et al., 2016, Mortuza et al., 2018, Singh et al., 2019a, Singh et al., 2020), there is still a need for potent and specific TLK inhibitors. Using the crystal structure of the kinase domain of TLK2 (Mortuza et al., 2018) proved useful for the virtual screening and modelling of compounds, but was insufficient for the rational design of specific TLK inhibitors. In Figures R8 and R9 we showed that several novel small molecules efficiently inhibited TLK1 and TLK2. These compounds could be used in cells and interacted synergistically with ATR, CHK1 and PARP inhibitors, as we have previously observed using siRNAs (Lee et al., 2018). However, the compounds are fairly non-specific and *in vitro* target another 60 kinases to differing extents. We repeated the *in vitro* kinase assays with a battery of derivates of both TLKi#1 and TLKi#3. Around half of the compounds inhibited TLKs while the others did not inhibit. These changes in efficacy could be explained by the modifications made on the original compounds and can be examined in detail when higher resolution structural details become available. Although the results presented in Figures R8 and R9 are a promising proof of principal, there is still a long road ahead to develop TLK inhibitors that are specific and potent enough to be tested *in vivo*.

Clinical implications of TLK2 mutations in neurodevelopment

In addition to its role in cancer, we examined the potential impact of TLK2 mutations in the etiology of MRD57, a poorly understood NDD. This included the molecular characterization of MRD57 variants, phenotyping of newly identified MRD57 patient cells and characterization of neurodevelopment in our TLK2 mouse model.

Functional characterization of MRD57 mutations was previously reported for only four variants in the PKD of TLK2 (TLK2- H493R, H518R, D629N, and R720*) (Figure I8). All of these mutations showed at least a 50% reduction in kinase activity as compared with TLK2-WT protein (Mortuza et al., 2018). In Chapter 3, we characterized 4 additional MRD57 mutations, two of which, TLK2-D551G and TLK2-S617L, are also located within the PKD of TLK2. TLK2-D551G and TLK2-S617L showed a profoundly impaired kinase activity, accompanied by altered subcellular localization, the significance of which remains unclear (Figures R19A-C, R20A-B and R21). TLK2-S617L caused a more severe impairment than TLK2-D551G, with *in silico* predictors also suggesting that TLK2-S617L was more likely pathogenic (Figure R19D). Interestingly, the patient carrying the p.(Asp617Leu) presented with a mild-to-severe ID, accompanied by a lower IQ than the patient carrying the p.(Asp551Gly) variant. This low IQ was comparable to that of the 4 patients that carried the p.(Glu475Ter) variant, that we thought would severely impair kinase activity as the truncation mutation is located in the middle of the PKD and results in the loss of a large part of the TLK2 protein (Figure R17F) (clinical details in (Pavinato et al., 2020)). This could suggest that the more severe impairment of TLK2 activity is reflected by a more severe clinical phenotype. This data further expands the characterization of MRD57 missense mutations and reinforces the prevailing hypothesis that MRD57 mutations located within the PKD cause a reduction of TLK2 kinase activity.

It therefore came as a surprise that the other 2 mutations we characterized in Chapter 3, TLK2-K55E and TLK2-G297D, are highly hyperactive and showed a normal pan-nuclear localization (Figures R20A, R20C and R21). Both mutations are located outside of the PKD with the TLK2-K55E mutation located very close to the predicted NLS (^{61}RNRKRKAEPY^{70}) starting at position

R61. We hypothesized that the mutation would impair nuclear localization which would in turn reduce activity by limiting access to ASF1, and potentially other substrates. In the case of TLK2-G297D, the mutation is located between the CC1 and CC2, and we thought it may affect the interaction of TLK2 with TLK1, which requires the CC1 domain (Mortuza et al., 2018). However, we demonstrated that is not the case (Figure R26). Both mutations are located in conserved amino acids (Reijnders et al., 2018, Topf et al., 2020), that are clearly important for activation but their precise roles are unknown. We also do not know if hyperactive mutations act as a monomer or maintain the need for dimerization to be active, as is the case of TLK2-WT (Mortuza et al., 2018).

How these mutations contribute to pathogenesis remains an open question. One possibility would be that both a reduction and an amplification of TLK activity impairs replication fork progression by impeding histone exchange. The precedent for this exists with ASF1 where it was shown that overexpression of either ASF1a phosphorylation deficient mutants or phospho-mimetic mutants impaired replication fork progression (Klimovskaia et al., 2014). This suggests a tightly regulated phosphorylation-dephosphorylation cycle that could be required for replication fork movement and indicates that a yet to be identified phosphatase activity must be intrinsically linked to TLK-ASF1 regulation. One candidate for this would be RIF1, a protein that functions in NHEJ, a process where ASF1A has also been implicated, as well as replication timing (Lee et al., 2017, Chapman et al., 2013, Escribano-Diaz et al., 2013, Alver et al., 2017). RIF1 has a conserved PP1 interaction site and we have identified it as a proximal interactor with TLK2 (Figure R27) (discussed in more detail below). However, we cannot rule out that hyperactivating mutations act through an entirely different patho-mechanism unrelated to replication fork progression or DNA repair.

Whether hyperactivation is a general feature of MRD57 mutations located outside of the PKD or is particular to these 2 residues, remains unclear. We also do not know if patients with hyperactivating mutations present with a distinct set of symptoms compared to patients with inactivating mutations due to the few patients identified with any particular mutation and lack of detailed clinical reporting in general. There is no clinical information available from the patient carrying the p.(Gly297Asp) mutation (Reijnders et al., 2018), and the p.(Lys55Glu) mutation is the only recessive mutation identified. The patient

carrying the p.(Lys55Glu) mutation presents with a more severe phenotype than previously reported in MRD57 patients, including severe motor and language delay, profound ID, microcephaly, cerebellar vermis hypoplasia and West syndrome (Topf et al., 2020). Whether the more severe phenotype is because the mutation is hyperactive or because it is in homozygosis is unknown. We favor the second possibility, as both parents are heterozygous for the variant and clinically unaffected. TLK2-K55E is the only TLK2 mutation described in homozygosis so far, highlighting that recessive variants in *TLK2* can also be disease causing. This dual role of some genes in which both loss-of-function and gain-of-function mutations or overexpression cause NDD has been previously described in other NDD-causing genes such as DNMT3A, FOXP1, and DYRK1A (Wegiel et al., 2008, Hamdan et al., 2010, Sollis et al., 2016, Duchon and Herault, 2016, Earl et al., 2017, Heyn et al., 2019, Lavery et al., 2020). For instance, gain-of-function mutations of the methyltransferase *Dnmt3a* cause microcephalic dwarfism due to DNA hypermethylation at developmental genes. On the other hand, loss of the methyltransferase *Dnmt3a* in GABAergic inhibitory neurons during development (but not in adulthood) leads to genome-wide hypomethylation, causing a misregulation of hundreds of genes and contributing to the phenotype of Rett Syndrome (Feng et al., 2010, Heyn et al., 2019, Lavery et al., 2020). This highlights the fact that the proper activity of several epigenetic regulators is critical for neurodevelopment.

The fact that almost all TLK2 mutations in MRD57 patients are *de novo* in heterozygosis, suggests that some level of wild type TLK2 expression is essential for development, and that homozygous deleterious mutations might be lethal in humans. Loss of TLK2 during embryonic development in mice is lethal due to a placental defect (Segura-Bayona et al., 2017). Despite this, we cannot directly extrapolate these findings into humans because of the marked differences in gestation length and brain development. We are not aware of data, but do not discard the possibility that mutations in TLK2 could cause placental issues or a higher risk of miscarriage. Related to this point, the pregnancy from family 1 (Figure R17B) started as bigeminal, with the twin of case 1 passing at week 9 of gestation. The rest of the pregnancy was considered high risk with several threats of abortion, and the baby suffering from oligoamnios and decreased fetal movements (Pavinato et al., 2020). A possibility would be that the p.(Asp551Gly) mutation caused defects in the placenta that led to a high risk pregnancy and the

death of the twin, although we do not know if the twins were identical and shared the placenta.

Despite the severe effects on kinase activity and clinical outcome caused by heterozygous missense mutations identified in MRD57 patients, mice heterozygous for *Tlk2* (*Tlk2*$^{+/-}$) are not embryonic lethal and their placentas are viable. Although *Tlk2*$^{+/-}$ mice did not seem to have any overt phenotype, neurodevelopment and behavior were not assessed (Segura-Bayona et al., 2017). We have yet to perform an in-depth behavioral evaluation of mice with reduced *Tlk2* levels, namely *Tlk2*$^{+/-}$ and *Tlk2*$^{-/-}$, to determine if they display features of ID or ASD and therefore represent a model of MRD57. There are several behavioral phenotyping tests that can be performed in mice that address both correct neurodevelopment, as well as the core symptoms of NDDs. These inlcude tests that assess cognition in the form of general sociability and interest in social novelty, ultrasonic vocalization of the pups towards the mother, and repetitive behaviors, such as self-grooming (Brown et al., 2000, Silverman et al., 2010, Premoli et al., 2019). If *Tlk2*$^{+/-}$ or *Tlk2*$^{-/-}$ mice show these behaviors, it could indicate that they are a relevant model for beginning to understand MRD57 etiology.

Implications of chromatin compaction in cells derived from MRD57 patients

We investigated several hypotheses based on our previous work that would potentially explain the role of TLK2 in neurodevelopment. We recently showed that total TLK depletion led to an inflammatory and innate immune response in cultured cancer cells. This response was triggered by the failure to maintain heterochromatin silencing that resulted in the spurious transcription of repetitive elements (i.e. telomeres, endogenous retroviruses (ERVs)) and induced features of alternative lengthening of telomeres (ALT), such as the production of extra-telomeric DNA (Segura-Bayona et al., 2020) (Figure I5C). There are numerous studies linking aberrant ERV expression with activation of the innate immune system and its contribution to the pathogenesis of inflammatory, neurodevelopmental and neurodegenerative diseases, indicating that this could

be relevant in MRD57 (Hurst and Magiorkinis, 2015, Saleh et al., 2019, Jonsson et al., 2020).

Consistent with the possibility that impaired chromatin maintenance and epigenetic dysregulation could underlie the pathological effects of reduced TLK2 activity in some MRD57 patients, we reported a more relaxed state of the chromatin in patient-derived LCLs carrying the p.(Asp551Gly) variant (Figures R18A and R18B) and fibroblasts carrying the 17q23.2 deletion (Pavinato et al., 2020). In the case of the p.(Asp551Gly) variant, this was accompanied by an increased susceptibility to DNA damage (Figure R18C and R18D; not tested for 17q23.2 deletion). We have previously observed that upon TLK depletion in cancer cells, nascent chromatin was more sensitive to micrococcal nuclease (MNase) digestion, proving a role of TLK activity in nucleosome assembly and maintenance of chromatin organization (Lee et al., 2018). However, with our current data we cannot conclude that this increase in chromatin accessibility leads to the desilencing of ERVs or other repetitive elements, or if these patients present with a generalized increase of pro-inflammatory cytokines, as has been described in patients with ASD (Masi et al., 2015). This could be addressed using additional methods, such as Assay for Transposase-Accessible Chromatin sequencing (ATAC-seq) and RNA-seq, to address what regions of the chromatin are becoming more accessible and the status of repetitive element and telomere silencing. Furthermore, examination of serum from MRD57 patients could help to determine the levels of circulating inflammatory molecules that may contribute to their pathology.

TLK2 is part of a network of ASD susceptibility proteins

In our TLK2 proximal interactome experiments, we consistently observed a considerable number of known proteins that when mutated confer susceptibility to developing NDDs (Figure R22). These proteins and many other proximal TLK2 interactors bind or modify histone variants and contribute to chromatin remodeling during DNA replication, transcription and repair. Some examples are JMJD1C (H3K9 demethylase), METTL3 (methyltransferase), HDAC5 (histone deacetylase), BRD4 (binds to acetylated histones), YEATS2 (chromatin reader), and CHD7 and CHD8 (chromatin remodelers). Mutations in TLK2, especially

inactivating mutations, reduced some of these proximal interactions (Figures R23 and R24). However, whether they have a functional relationship relevant to MRD57 remains to be determined.

Several computational and experimental studies of protein-protein interaction (PPI) networks showed an overall excess in interactions between ASD-susceptibility proteins, showing they are closely connected within PPI networks. In particular, pathways identified by these studies included chromatin remodeling, transcriptional and translational regulation (Sakai et al., 2011, Corominas et al., 2014, Pinto et al., 2014, Satterstrom et al., 2020). These enhanced connections carry an important consequence: a point mutation in a given gene (say TLK2) has the potential to alter the function of other ASD genes, and can thus spread through the protein network. This could result in indirect changes in the protein levels and phosphorylation status of other NDD-susceptibility proteins that are not necessarily a direct interactor or substrate of TLK2.

From all the proximal TLK2 interactors, of particular interest to us are the ATP-dependent Chromodomain Helicase DNA Binding Proteins (CHD) CHD7 and CHD8. CHD proteins play an important role in chromatin remodeling during neurodevelopment, and thus regulate the transcription levels of developmentally-regulated genes such as those implicated in synaptogenesis. In the case of CDH8, its targets include many genes that have also been implicated in ASD (Cotney et al., 2015, Feng et al., 2017b, Manning and Yusufzai, 2017, Marie et al., 2018). Loss-of-function mutations in CHD7 and CHD8 are associated with neurodevelopmental defects: CHD7 haploinsufficiency causes CHARGE syndrome (OMIM 214800) (Lalani et al., 2006), while CHD8 is one of the strongest ASD high-risk-associated genes (OMIM 615032) (Bernier et al., 2014). Both proteins are enriched at replication forks (Wessel et al., 2019) and are known to interact with each other (Batsukh et al., 2012, Feng et al., 2017a) and with TLK2 (Figure R26). Phosphorylation in specific residues of CHD8 and other TLK2 proximal interactors (DIDO1 and METTL3) are lost upon TLK1+TLK2 silencing (Supplementary figure 1), suggesting that these sites are potentially regulated by TLKs either via direct phosphorylation or indirectly by another kinase regulated by TLKs. In future work, it will be important to examine the potential functional relevance of these interactions, as it is tempting to speculate that alterations in the interaction between TLK2 and the CHD proteins could be

contributing to the altered chromatin state observed in LCLs from the patients carrying the p.(Asp551Gly) variant.

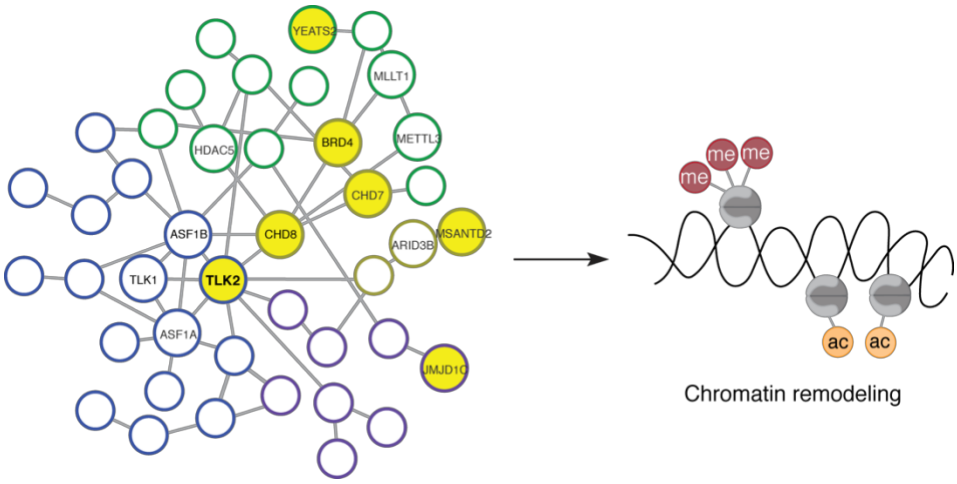


Figure D1. Network graph representing a simplified scheme of the TLK2 interactome.

ASD-susceptibility proteins are closely connected within protein-protein interaction (PPI) networks, especially in networks related to chromatin remodeling, transcriptional and translational regulation. This graph shows how a mutation in an ASD-susceptibility gene, namely TLK2, can affect a PPI network that includes numerous chromatin remodelers. Lines indicate PPIs and yellow nodes indicate known ASD susceptibility proteins found in the SFARI or DECIPHER databases.

TLK2 in histone regulation and inflammation in the mouse embryonic brain

When we examined gene expression in the telencephalon of *Tlk2*$^{-/-}$ Sox2Cre^{+} mice, we could readily observe that H3.3 (encoded by H3f3b) was the most downregulated gene, followed by other histone variants such as H2A, H2B and the linker histone H1 (encoded by Hist3h2a, Hist3h2ba, Hist1h1c, H1fx) (Figures R13B and R13C), although this was not the case in *Tlk2*$^{-/-}$ placentas (Figure R10E). Canonical histones (including H2A, H2B, H3.1/2, and H4) are encoded by multiple gene copies which are exclusively expressed during S phase, to meet the high demand of histones needed for chromatin compaction during DNA replication (Reviewed in Günesdogan et al., 2014). It is well established that an

interruption of the S phase leads to a rapid degradation of histone transcripts, and thus a reduction of histone transcription could reflect a reduction in proliferating cells in the telencephalon of *Tlk2*$^{-/-}$ Sox2Cre^{+} mice.

Furthermore, we have previously observed a decrease in *de novo* H3.3 histone deposition (Lee et al., 2018) and reduced H3.3 and H3K9me3 occupancy in telomeres and other genomic regions (Segura-Bayona et al., 2020) upon TLK depletion in cancer cells. H3.3 is a replacement histone variant required for the establishment of H3K9me3 marks to supports heterochromatin regions and the silencing of ERVs (Udugama et al., 2015, Elsasser et al., 2015), being essential for the maintenance of genome integrity during mammalian development (Jang et al., 2015). Postnatally, H3.3 completely replaces the canonical histone variants H3.1 and H3.2 in the mouse and human brains (Maze et al., 2015). Germline mutations in H3.3 have recently been associated with a distinct neurodegenerative disorder (Maver et al., 2019, Bryant et al., 2020), and mutations in other genes involved in histone regulation and epigenetic maintenance have been identified in patients with NDDs. Some examples include *ATRX*, a gene that regulates H3.3 deposition in heterochromatin (Udugama et al., 2015, Sadic et al., 2015), that is mutated in patients with ID and most ALT+ cancers (Moncini et al., 2013); and HIRA, a histone chaperone that deposits and recycles H3.3 during transcription (Torné et al., 2020), whose haploinsufficiency causes defects in neurodevelopment (Jeanne et al., 2021). In mice, *in utero* silencing of HIRA induces terminal mitosis and cell cycle exit, thus reducing NECs proliferation in the VZ, SVZ and IZ and resulting in premature neuronal differentiation (Li and Jiao, 2017). Furthermore, prenatal exposure to the histone deacetylase inhibitor valproic acid (VPA) led to increased expression of ERVs and has been linked to ASD in children and to behavioral alterations in mice (Kataoka et al., 2013, Christensen et al., 2013). All of these examples highlight the importance of accurate regulation of histone levels and histone post-translational modifications in the developing nervous system.

Apart from downregulation of histone variants, we observed an upregulation in the transcription of inflammatory, IFN-related, and NF-kB genes in the telencephalon of our E18.5 *Tlk2*$^{-/-}$ Sox2Cre^{+} mice (Figures R13B and R14A). This response was dependent on the dosage of *Tlk2*, as *Tlk2*$^{+/-}$ Sox2Cre^{+} mice presented with a milder, but still significant, phenotype (Supplementary figure 2). We also noted an upregulation of genes that corresponded to brain stromal cells, including microglia, endothelial cells, and a subtype of astrocytes (Aldoc^{high}), as

well as a downregulation of the cortical neurons signature (Figures R15A and R15B). It is important to remember that gliogenesis starts at around this time in development (E18 -ish), and that is when RGCs generate macroglial cells (astrocytes and oligodendrocytes) (Figure I2) (Gotz and Barde, 2005). That is why it does not come as a surprise that there is a notable overlap in the genes expressed by Aldoc^{high} astrocytes and two subtypes of precursor RGCs (Id3^{high} and gdf10^{high}), with both Id3 and gdf10 being astrocytic lineage markers. In the case of microglia, however, there is no overlap with genes expressed in other brain cell types, as they have a completely different embryonic origin (Supplementary table 4). Upregulation of immune gene modules (including IFN and NF-kB pathways) and glial markers (microglia and astrocytes) have been observed in seminal transcriptome-wide association studies of postmortem brains of idiopathic ASD, as well as other major psychiatric disorders. A downregulation of neuronal and synaptic modules was also observed in ASD and schizophrenia patients (Voineagu et al., 2011, Gupta et al., 2014, Gandal et al., 2018b). Another important observation is that the balance in NECs proliferation, differentiation and maturation is altered in ASD patients (Parikshak et al., 2013).

A cause of the increase in the astrocyte signature could be a premature start of gliogenesis. This has been previously reported after conditional deletion of the DNA methyltransferase I (*DNMT1*) in NECs. *Dnmt1* loss during embryogenesis leads to a genome-wide DNA hypomethylation in the NECs, including the demethylation of astrocytic markers and genes of the gliogenic JAK-STAT pathway, which contribute to an early astrocyte differentiation (Fan et al., 2005). In the context of glioma, loss of *ATRX* in a *P53*-null background drives a global change in chromatin accessibility, histone composition, and transcription, shifting the differentiation of the NECs towards the astrocytic lineage (Danussi et al., 2018). This suggests that chromatin remodeling in the NECs could lead to increased chromatin accessibility and transcription of key genes associated with astrocytic lineage in the context of RGC differentiation, such as *Id3* and *gfap*, both of which we observe an increased transcription in the telencephalon of *Tlk2*$^{-/-}$ Sox2Cre^{+} mice (p-v = 0.005 and 0.02, respectively, in our dataset *Tlk2*$^{-/-}$ Sox2Cre^{+} vs *Tlk2*$^{+/+}$ Sox2Cre^{+}).

The source of inflammation in the brain of the *Tlk2*$^{-/-}$ Sox2Cre^{+} mice embryos remains to be identified. A recent study showed that the expression of high levels of ERVs during murine brain development in the proliferating NECs, but not in

mature neurons in the adult, causes an inflammatory response that is associated with activated microglia and aggregates of ERV-derived proteins (Jonsson et al., 2021). We hypothesized that this upregulation of inflammatory and IFN-related gene transcription in the brain could come from the spurious transcription of repetitive elements such as ERVs, as we have previously seen in cultured cancer cells (Segura-Bayona et al., 2020). However, analysis of our telenchephalon RNAseq data using Repeat Masker (Bao et al., 2015) did not reveal a clear upregulation of repetitive elements in the different repetitive element classes (Supplementary figure 3). This could mean that the effects are unrelated to ERVs expression or that we do not observe an increase in ERVs due to technical reasons. For both our brain and placenta RNA-seq, we enriched for polyA$^+$ mRNAs, and despite many studies looking at repetitive elements on polyA$^+$-enriched RNA-seq, all the non-coding RNAs are not in the mix and RNA repeats are underrepresented (Hrdlickova et al., 2017, Solovyov et al., 2018). Another factor could be that we only knocked out TLK2 in the mice, while in U-2-OS cells only the depletion of both TLK1 and TLK2 caused a marked increase in transcription of ERVs and other repetitive elements (Segura-Bayona et al., 2020). Another caveat of our genetic mouse model is that the TLK2 knockout is ubiquitous and affects every cell type of the mice, while in (Jonsson et al., 2021) they use a conditional Cre that is expressed only in cortical progenitors starting from E10.

Another open question is what drives the increase in the transcriptional signature of microglia in the telencephalon of *Tlk2*$^{-/-}$ Sox2Cre$^+$ and *Tlk2*$^{+/-}$ Sox2Cre$^+$ mice (Figure R15). It is known that inflammation caused by LPS injection in a MIA model causes an increase in the number of microglia in the developing cortex. A possible option would be that the *Tlk2*$^{+/-}$ mothers presented with systemic inflammation, although we have reasons to believe that is not the case. Assessment of the levels of IL-6, the cytokine that drives MIA (Hsiao and Patterson, 2011) in non-pregnant adult *Tlk2*$^{+/-}$ mice by ELISA did not show increased levels of IL-6 as compared to *Tlk2*$^{+/+}$ mice (not shown). Moreover, the increase in microglia in MIA models has been associated to the reduction of Pax6$^+$ and Tbr2$^+$ progenitors that results in a reduction in thickness of the VZ/SVZ areas. Microglia has been shown to control the number of progenitors and neurons in the developing cortex by phagocytosing not only apoptotic cells but also proliferating progenitors and mature neurons (Cunningham et al., 2013, Tronnes et al., 2016, Tong and Vidyadaran, 2016). In the contrary, mice lacking

microglia such as *Csf1r*$^{-/-}$ mice display brain abnormalities such as increased brain mass due to an expanded number of NECs (Erblich et al., 2011, Nandi et al., 2012). This suggests that the reduction in cortical neurons signature (Figure R15) could be in part due to increased phagocytosis by microglial cells.

TLK2's proximal interactome and potential roles of TLKs in DNA damage repair and replication timing

TLK activity is highest during S phase when there is a high need for histone deposition and is attenuated upon RS by CHK1 direct phosphorylation of TLK1 (Figure I3) (Silljé et al., 1999, Groth et al., 2003, Krause et al., 2003). Although TLK2 has similar phospho sites at its C-terminus, it remains unclear if TLK2 is also directly phosphorylated by CHK1. A possibility would be that TLK2 inhibition happens through its heterodimerization with TLK1, as TLK1 (but not TLK2) was identified as a direct substrate of CHK1 using an analog-sensitive allele (Blasius et al., 2011). However, we have not identified CHK1 in our proteomics experiments (Figures R22D, Supplementary figure 4), and the direct physical interaction between TLK1 and CHK1 is not well characterized, although TLK1 was identified as a CHK1 substrate using an analog sensitive allele of CHK1 (Blasius et al., 2011). One possibility would be that the interaction is too transient to be detected by IP-MS or BioID. Another possibility that we will discuss later in further detail would be that we do not see CHK1 in our experiments because we performed them in asynchronous cells and in the absence of DNA damage. Overall, the mechanisms of specific spatio-temporal regulation of TLKs during the cell cycle remains unclear.

Both TLK1 and TLK2 have been consistently identified in isolation of proteins on nascent DNA (iPOND) studies as proteins associated with replication forks and nascent chromatin (Sirbu et al., 2011, Wessel et al., 2019), and were also identified in a PCNA-BioID experiment in synchronized S phase cells (Srivastava et al., 2018). Our extensive MS studies (Figures R22D, Supplementary figure 4 and (Segura-Bayona et al., 2017)) show how the proximal interactome of TLK2 contains numerous proteins enriched at replication forks. Some of these proximal interactors have roles in DNA repair such as RAD50 and MutS Homolog 3 (MSH3), transcriptional repression such as NACC1 and SCML2, and chromatin remodeling such as CHD7 and CHD8 (Figures R22D and R26). While this does

not necessarily mean these proteins are direct interactors, it provides insight into the cellular environment of the TLKs, and is consistent with the idea that TLKs provide histones to replication via interaction and phosphorylation of ASF1 (Klimovskaia et al., 2014, Lee et al., 2018). We do not know if TLK1 and TLK2 always have the same localization pattern, if they are always at replication forks or by the contrary they are recruited there when needed, potentially serving distinct functions. In future work, we will examine which interactions are necessary for their localization at forks and the consequences of perturbing them.

Apart from the already known and validated interactors of TLK2 (ASF1a, ASF1b, and LC8 type I and II (DYNLL1 and DYNLL2)), another reproducible proximal interactor of TLK2 that we identified is replication timing regulatory factor 1 (RIF1). RIF1 has come out in in all our MS experiments, including IP-MS and BioID (Segura-Bayona et al., 2017), although below our stringent statistical threshold due to the fact that it is labeled in some control experiments. Even though RIF1 was first identified in yeast as a regulator of telomere length (Shi et al., 2013), this telomeric function is not maintained in mammals. However, mammalian RIF1 is associated with heterochromatin and is implicated in the control of replication timing and nuclear architecture (Cornacchia et al., 2012, Foti et al., 2016, Alver et al., 2017), as well as in DSB repair via the NHEJ pathway (Chapman et al., 2013, Escribano-Diaz et al., 2013). We have been able to validate the interaction between TLK2 and RIF1 by performing IP of overexpressed TLK2 and blotting for endogenous RIF1 (Figure R27A). However, this experimental approach does not allow us to assess if the interaction is direct or not. Published IP-MS analysis of mouse embryonic stem cells (mESC) on RIF1-associated complexes identified TLK1 and TLK2 (Sukackaite et al., 2017), while similar experiments performed in ASF1a in cancer cells identified RIF1 (Yang et al., 2018), suggesting that the RIF1-TLK interaction could occur through ASF1a. We have validated that RIF1 interacts with both ASF1a and ASF1b (Figure R27B), but we have been unable to decipher if the interaction between TLK2-RIF1 is direct or through ASF1s.

The way in which RIF1 controls replication timing is by interacting with and serving as a substrate-targeting scaffold for protein phosphatase 1 (PP1). RIF1 directs PP1 to licensed replication origins and promotes PP1-mediated dephosphorylation of MCM4, counteracting the DDK-dependent phosphorylation necessary for the activation of the MCM complex and

the initiation of DNA replication, therefore keeping late origins dormant (Hiraga et al., 2014, Alver et al., 2017). As RIF1 acts partly through the recruitment of a phosphatase, it is tempting to speculate that it can exert some kind of regulatory effect on TLKs. RIF1-PP1 may act upstream of TLKs as a direct regulator, by dephosphorylating TLKs autophosphorylation sites and turning them off, or by counteracting TLK phosphorylation of ASF1s or other unidentified substrates in a manner similar to DDK. We recently used ATAC-seq to examine TLK-depleted U-2-OS cells and observed a decrease in accessibility of early replication origins and an increase in accessibility of late replication origins upon TLK loss (Segura-Bayona et al., 2020). This suggests that TLKs could contribute to maintaining dormant late origins and replication timing, possibly through RIF1. In mouse ES cells, RIF1 coats late-replicating domains protecting them from premature firing (Foti et al., 2016). Loss of TLKs could potentially cause a reduction of RIF1-PP1 localization at replication forks, maintaining the phosphorylation of MCM4 and therefore causing premature late-origin firing. In future work we will determine if RIF1 is a substrate or regulator of TLKs and address these possibilities.

Human TLK1 has been proposed to have a role in DNA damage repair, via regulation of RAD9 and NEK1. Although TLK1 has been previously shown to interact and phosphorylate both proteins (Singh et al., 2017b, Sunavala-Dossabhoy and De Benedetti, 2009, Kelly and Davey, 2013), we have not identified NEK1 or RAD9 interactions in IP-MS or BioID experiments (Figure R22D, Supplementary figure 4, (Segura-Bayona et al., 2017, Pavinato et al., 2020). Furthermore, in IP experiments RAD9 was pulled down by both TLK1 and TLK2 at similar levels, although it was equally enriched in empty vector control samples in the presence and absence of benzonase, suggesting that RAD9 is a potential contaminant (Segura-Bayona, 2018). However, we cannot discard that this interaction happens under DNA damaging conditions or is markedly different in other cell types.

Three TLK interactors that we have validated, ASF1a, LC8 and RIF1, have all been implicated in DSB repair via the NHEJ pathway (Chapman et al., 2013, Escribano-Diaz et al., 2013, Lee et al., 2017, Becker et al., 2018). TLK1 and TLK2, together with RIF1 were identified as proximal interactors of 53BP1 by APEX2 labeling (Gupta et al., 2018), while TLK2 was identified as a proximal interactor of MDC1, which interacts with ASF1a (Lee et al., 2017, Gupta et al.,

2018). ASF1a has been proposed to function in DSBs repair independent of its histone chaperone activity (Lee et al., 2017). ASF1a is phosphorylated by the DNA-PKcs kinase that is activated by DSBs and interacts with MDC1 to facilitates its interaction with activated ATM at the site of damage. This starts a cascade of events that include the phosphorylation of RNF8-RNF168 and the ubiquitination of gH2AX, that facilitate the recruitment of 53BP1-RIF1 complexes (Lee et al., 2017). Notably, RNF8 is also a consistent proximal interactor of TLK2 but we have not investigated this further (Figure R22D). 53BP1 binds ubiquitinated gH2AX and is the phosphorylated by ATM. RIF1 is then recruited to the site of damage by direct interaction with 53BP1. 53BP1-RIF1 then act together to promote NHEJ by preventing BRCA1-CtIP complexes from binding DSB sites and mediating processing of the breaks that precedes HR (Chapman et al., 2013, Escribano-Diaz et al., 2013). LC8 localizes at the sites of damage by association with 53BP1. Loss of LC8, 53BP1, RIF1 and some of their interactors disrupts DSB repair via NHEJ, restores DNA end resection and HR, conferring resistance to the synthetic lethality of PARP inhibition in HR-deficient cells (He et al., 2018, West et al., 2019). Despite the fact that both TLKs and ASF1s interact with RIF1 (Figure R27A and R27B), we have been unable to clearly link TLK activity with NHEJ (Segura-Bayona, 2018). By addressing DSB repair pathway choice (Bennardo et al., 2008), we observed a decrease in NHEJ upon TLK2 silencing but to a lesser extent than when silencing ASF1a. However, this needs to be further studied as we did not see an increase in HR upon TLK2 silencing, nor could we fully recapitulate the increase in HR upon ASF1a silencing as previously proposed (Lee et al., 2017).

When interpreting our BioID results, it is important to note that we did the experiments in asynchronous cells. BioID has a slow kinetics and 24 hour biotin labelling are needed to obtain enough biotinylated material for proteomic analysis (Roux et al., 2012). A new version of BirA*, TurboID, allows for a faster biotinylation, obtaining similar levels than BioID in only 10 minutes (Branon et al., 2018). Another advantage of TurboID is that it allows for organelle specific localization, as it is coupled to, for example, an NLS. Given that TLKs have nuclear expression and their highest activity is during S phase, it would be interesting to study the interactome in synchronized cells using TurboID. This approach would also allow us to study the interactome in response to several DNA damaging insults, and help us elucidate if TLKs are involved in DSB repair

or sensing, as their proximal interactors RIF1, ASF1a, and the MRN complex (Figure R26).

While many substrates have been proposed, only ASF1 has looked like a bona fide substrate. However, we believe it is unlikely that ASF1 is the only substrate of TLKs, although so far, we have not been unable to identify other substrates. We have attempted to identify TLK2 substrates using an analog-sensitive (AS) kinase approach (Segura-Bayona, 2018). However, we did not manage to generate a TLK2 AS kinase that was sensitive to the bulky ATP analogs. We next performed phospho-proteomics experiments in the presence and absence of TLK1+2 in order to try to identify what phosphor sites are lost upon TLK loss (Supplementary figure 1). Although we have obtained some interesting hits (discussed above), the analysis has proven very difficult due to data complexity, lack of a clear TLK consensus sequence and the fact that we did the experiment on whole protein lysates. In future work, we will refine our approaches to try and identify the substrate spectrum of TLKs and determine their functional relevance in cancer therapy and development.

Conclusions

- TLK2 expression levels negatively correlate with the expression of markers of immune system infiltration in several tumors types from TCGA datasets.
- Knocking out *Tlk2* in already formed mammary tumors stalls primary tumor growth and strongly reduces lung metastasis incidence.
- *Tlk2* knockout tumors exhibit increased immune infiltration at early stages.
- TLK is a potential druggable target for cancer therapy.
- Novel small molecules TLKi#1 and TLKi#3 inhibit TLK activity.
- *Tlk2* knockout placentas do not exhibit increased inflammation.
- *Tlk2* knockout telencephalons have normal alternative splicing.
- *Tlk2* knockout telencephalons exhibit elevated expression of inflammatory genes.
- *Tlk2* knockout telencephalons exhibit upregulation of stromal cell components, including microglia and astrocytes.
- TLK2 mutations disrupt proper chromatin compaction in patient-derived lymphoblastoid cells.
- The proximal interactome of TLK2 includes many proteins known to be involved in autism and other neurodevelopmental disorders.
- The proximal interactome of TLK2 includes many proteins found at active replication forks, including components of the MRN complex.
- TLK2 missense mutations alter the proximal interactome of TLK2.

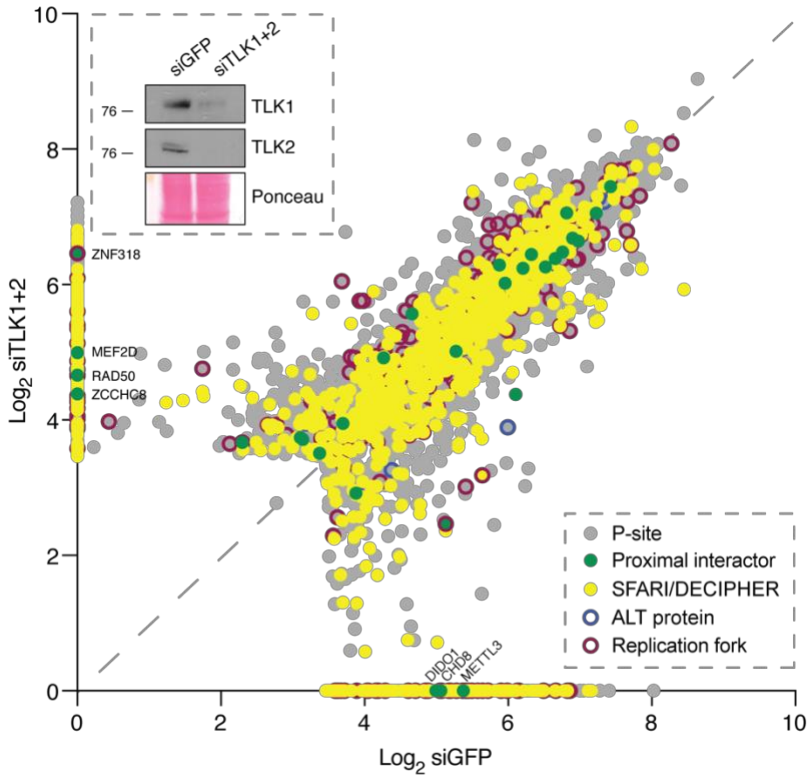
- TLK2-D551G and TLK2-S617L missense mutations impair kinase activity and normal subcellular localization.

- The TLK2-K55E and TLK2-G297D mutations lead to hyperactive kinase activity.

Supplementary information

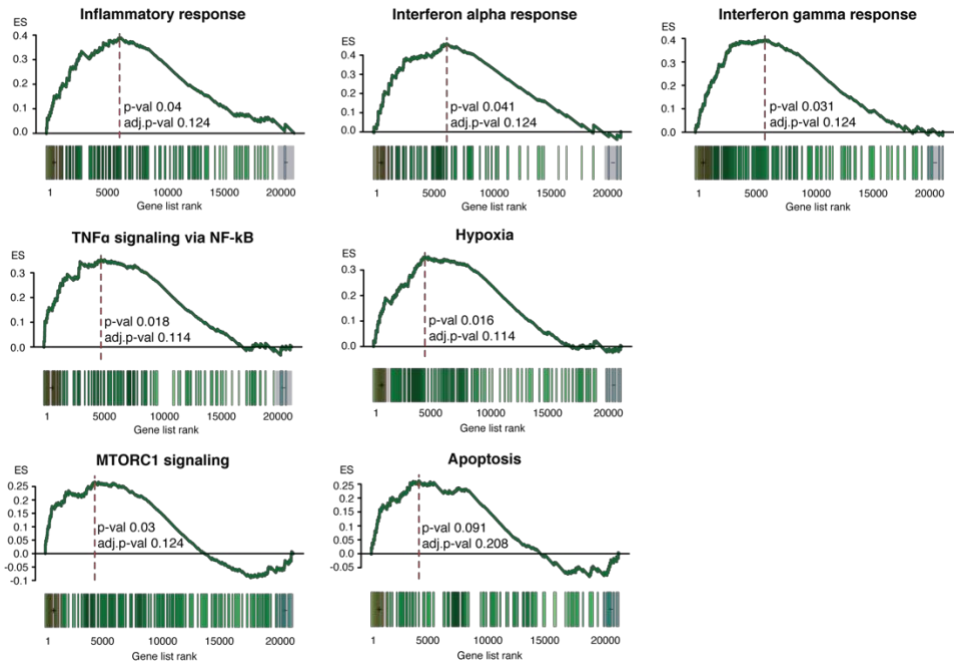
Supplementary figure 1. Phospho-proteomics analysis of hTERT-RPE-1 cells.

Scatter plot of average spectral counts (Log_2 transformed) of phospho-sites identified in hTERT-RPE-1 control cells (siGFP) compared to hTERT-RPE-1 cells depleted of TLK1 and TLK2 for 48 hours (siTLK1+2). Each dot represents a phospho-site in a given protein. Phospho-sites in proteins that are listed in the SFARI/DECIPHER databases are shown in yellow, while proximal interactors of TLK2 (determined by the BioID experiments shown in Figure R22) are shown in green. Blue circles indicate phosphor-sites in alternative lengthening of telomeres (ALT) proteins, while purple circles indicate proteins enriched in replication forks (Wessel et al., 2019). The western blot shows proper depletion of TLK1 and TLK2. Ponceau staining shows equal loading. The experiment was done once.



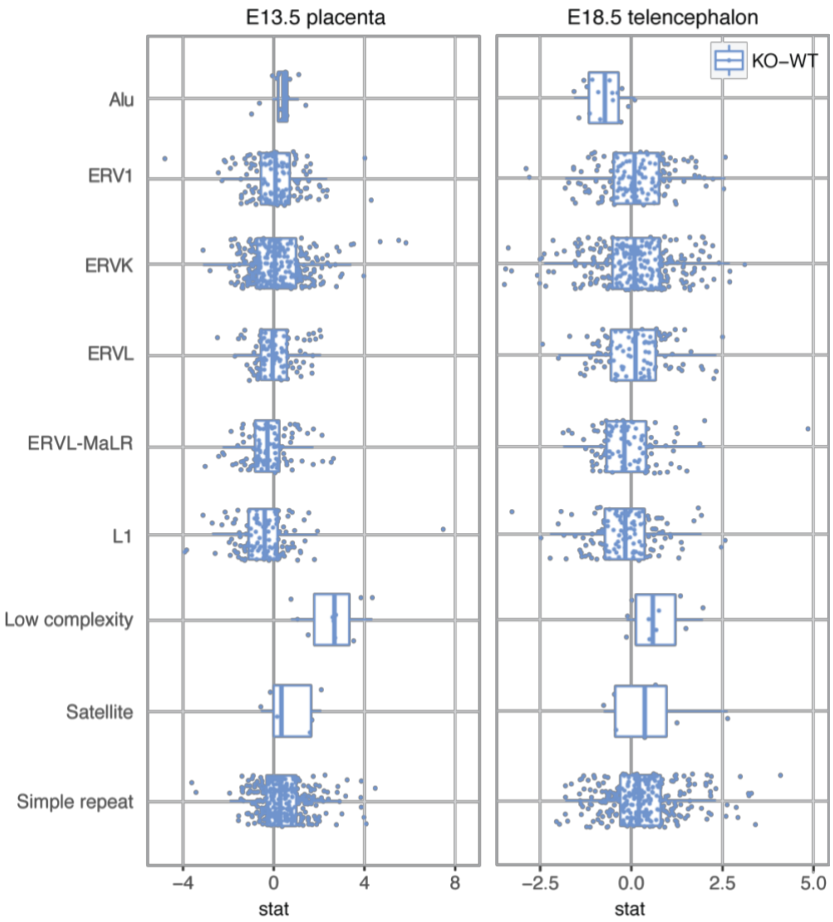
Supplementary figure 2. Increased expression of neuroinflammatory genes in TLK2 heterozygous embryos.

GSEA where genes were annotated to Hallmark terms (Liberzon et al., 2015). Some of the categories upregulated in *Tlk2*$^{+/-}$ Sox2-Cre^{+} telencephalons as compared to *Tlk2*$^{+/+}$ Sox2-Cre^{+} are shown. This data corresponds to that in Figures R11-R15.



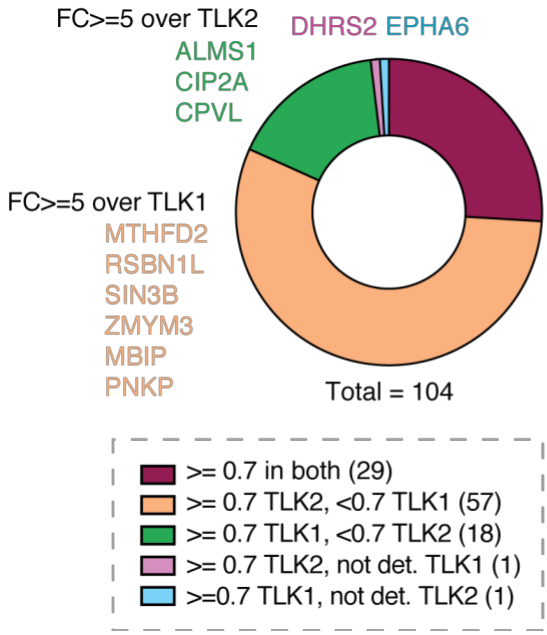
Supplementary figure 3. Repetitive element analysis of E13.5 placentas and E18.5 telencephalons.

Annotations for repetitive elements were obtained from the repeatmasker website (Smit, 2013-2015). Counts per repetitive element were obtained using the Rsubread package (Liao et al., 2013). DESeq2 (Love et al., 2014) was considered to evaluate the differences between WT and KO samples. Every dot represents the average counts per repetitive element of TLK2 KO – WT. Selected repetitive element classes are shown. N=3 male placentas/genotype and n=3 male and n=3 female telencephalons/genotype were sequenced. Placenta data corresponds to that in Figure R10 while telencephalon data corresponds to that in Figures R11-R15.



Supplementary figure 4. Comparison of the proximal interactome of TLK1 and TLK2.

Data obtained from BioID experiments. Confidence interactors with SAINT>=0.7 in both TLK1 and TLK2 are shown in burgundy. FC, fold change.



ASF1A
ASF1B
CHD7
YEATS2
HIRA
MSANTD2
NELFCD
GPATCH1
XAB2
NFAT5
UBN1
HDAC5
POLH
CACTIN
NCOA6
TPR
RAD50
ZNF318
DYNLL1
NACC1
DYNLL2
ZBTB33
ZHX3
ZNF131
KNL1
MYBL2
WDR83
METTL3

Supplementary table 1. TCGA cohort designation and number of patients per cohort (n).

ER, estrogen receptor; HER2, human epidermal growth factor receptor 2.

Abbreviation	Cancer Type	n
ACC	Adrenocortical carcinoma	75
BLCA	Bladder urothelial carcinoma	397
BLCA.nonpap	Bladder urothelial carcinoma. Non-papillary	260
BLCA.pap	Bladder urothelial carcinoma. Papillary	132
BRCA	Breast invasive carcinoma	1033
BRCA.basal	Breast invasive carcinoma. Basal-like	131
BRCA.erneg	Breast invasive carcinoma. ER negative	239
BRCA.erneg.her2neg	Breast invasive carcinoma. ER negative, HER2 negative	193
BRCA.erneg.her2pos	Breast invasive carcinoma. ER negative, HER2 positive	46
BRCA.erpos	Breast invasive carcinoma. ER positive	794
BRCA.erpos.her2neg	Breast invasive carcinoma. ER positive, HER2 negative	740
BRCA.erpos.her2pos	Breast invasive carcinoma. ER positive, HER2 positive	54
BRCA.her2	Breast invasive carcinoma. HER2	64
BRCA.her2neg	Breast invasive carcinoma. HER2 negative	933
BRCA.her2pos	Breast invasive carcinoma. HER2 positive	100
BRCA.luma	Breast invasive carcinoma. Luminal A	405
BRCA.lumb	Breast invasive carcinoma. Luminal B	170
BRCA.normal	Breast invasive carcinoma. Normal-like	20
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma	293
CHOL	Cholangiocarcinoma	36
COAD	Colon adenocarcinoma	428
COAD.msi	Colon adenocarcinoma microsatellite instable	77
COAD.mss	Colon adenocarcinoma microsatellite stable	348
DLBC	Lymphoid neoplasm diffuse large B-cell lymphoma	47
ESCA	Esophageal carcinoma	162
GBM	Glioblastoma multiforme	155

Abbreviation	Cancer Type	n
HNSC	Head and neck squamous cell carcinoma	502
KICH	Kidney chromophobe	64
KIRC	Kidney renal clear cell carcinoma	481
KIRP	Kidney renal papillary cell carcinoma	280
KIRP.type1	Kidney renal papillary cell carcinoma.Type1	73
KIRP.type2	Kidney renal papillary cell carcinoma.Type2	84
LAML	Acute myeloid leukemia	108
LGG	Brain lower grade glioma	516
LIHC	Liver hepatocellular carcinoma	355
LUAD	Lung adenocarcinoma	499
LUSC	Lung squamous cell carcinoma	479
MESO	Mesothelioma	81
OV	Ovarian serous cystadenocarcinoma	289
PAAD	Pancreatic adenocarcinoma	159
PCPG	Pheochromocytoma and paraganglioma	164
PRAD	Prostate adenocarcinoma	472
READ	Rectum adenocarcinoma	153
SARC	Sarcoma	246
SKCM	Skin cutaneous melanoma	459
STAD	Stomach adenocarcinoma	401
STAD.msi	Stomach adenocarcinoma microsatellite instable	77
STAD.mss	Stomach adenocarcinoma microsatellite stable	322
TGCT	Testicular germ cell tumors	138
THCA	Thyroid carcinoma	469
THYM	Thyoma	102
UCEC	Uterine Corpus Endometrial Carcinoma	514
UCS	Uterine Carcinosarcoma	56
UVM	Uveal melanoma	79

Supplementary table 2. Comparison between TLK2 transcripts NM_001284333.2 and NM_006852.6.

In this thesis, all variants are referred to GRCh37 annotation and to NM_001284333.2, in line with the previously published TLK2 crystal structure (Mortuza et al., 2018). For homogeneity with the clinical work from (Reijnders et al., 2018), we specify here variants also in NM_006852.6.

NM_001284333.2 (Q86UE8-1; isoform 1)		**NM_006852.6 (Q86UE8-2; isoform 2)**	
c.163A>G	p.(Lys55Glu)	c.163A>G	p.(Lys55Glu)
c.890G>A	p.(Gly297Asp)	c.890G>A	p.(Gly297Asp)
c.1423G>T	p.(Glu475Ter)	c.1357G>T	p.(Glu453Ter)
c.1652A>G	p.(Asp551Gly)	c.1586A>G	p.(Asp529Gly)
c.1850C>T;	p.(Ser617Leu)	c.1784C>T	p.(Ser595Leu)

Supplementary table 3. List of SFARI /DECIPHER /OMIM genes in the proximal interactome of wild-type and mutated TLK2.

Gene	SFARI score	Disease	OMIM ID	Inheritance
ALX4		{Craniosynostosis 5, susceptibility to}	615529	AD
		Frontonasal dysplasia 2	613451	AR
		Parietal foramina 2	609597	AD
BARD1		breast cancer, susceptibility to	114480	AD, Smu
BRD4	3		608749	
BRIP1		{Breast cancer, early-onset, susceptibility to}	114480	AD, SMu
		Fanconi anemia, complementation group J	609054	
CCNK	S	?Intellectual developmental disorder with hypertelorism and distinctive facies	618147	
CHD7	1	CHARGE syndrome	214800	
	1	Hypogonadotropic hypogonadism 5 with or without anosmia	612370	
CHD8	1	{Autism, susceptibility to, 18}	615032	AD
CIC	1	Mental retardation, autosomal dominant 45	617600	AD
EBF3		Hypotonia, ataxia, and delayed development syndrome	617330	AD
FLI1		Bleeding disorder, platelet-type, 21	617443	AD, AR
FOXC1				
GATAD1		?cardiomyopathy, dilated, 2B	614672	AR
JMJD1C	3		604503	
KIF4A		?Mental retardation, X-linked 100	300923	XLR
MCM8		?premature ovarian failure 10	612885	AR
MLH1		Colorectal cancer, hereditary nonpolyposis, type 2	609310	
		Mismatch repair cancer syndrome	276300	AR
		Muir-Torre syndrome	158320	AD
MSANTD2	3			
MSH3		Endometrial carcinoma, somatic	608089	
		Familial adenomatous polyposis 4	617100	AR

Gene	SFARI score	Disease	OMIM ID	Inheritance
MYSM1		Bone marrow failure syndrome 4	618116	AR
NACC1	1	Neurodevelopmental disorder with epilepsy, cataracts, feeding difficulties, and delayed brain myelination	617393	AR
NFIA	3	Brain malformations with or without urinary tract defects	613735	
NFIX	S	Marshall-Smith syndrome	602535	AD
		Sotos syndrome 2	614753	AD
NSD1	1	Sotos syndrome 1	117550	AD
OGT				
PAPOLG	3			
PMS2		Colorectal cancer, hereditary nonpolyposis, type 4	614337	
		Mismatch repair cancer syndrome	276300	AR
PNKP				
PRPF3		Retinitis pigmentosa 18	601414	AD
RAD50		Nijmegen breakage syndrome-like disorder	613078	
RBM10		TARP syndrome	311900	XLR
THOC2		Mental retardation, X-linked 12/35	300957	XLR
WRN		Werner syndrome	277700	AR
YEATS2	3	?Epilepsy, myoclonic, familial adult, 4	615127	
ZCCHC8		?Pulmonary fibrosis and/or bone marrow failure, telomere-related, 5	618674	AD
ZNF148		Global developmental delay, absent or hypoplastic corpus callosum, and dysmorphic facies	617260	AD

SFARI (release 2020Q1) / DECIPHER (v.9.32)

OMIM

AD=Autosomal Dominant; AR=Autosomal Recessive; SMu= Somatic Mutation; XLR=X-linked recessive

Significance defined as SAINT > or = 0.7 in at least one condition (TLK2-WT or mutants)

Supplementary table 4. Number of overlapping genes between the different Mouse Cell Atlas cell populations.

Geneset	Macrophage/Microglia	Ependymal cells	GABAergic neuron_Nrxn3^{high}	Cortical neurons	GABA-Interneurons	Erythrocyte	Dopaminergic amacrine neuron	Immature neuron	Glutamatergic neuron	Hypothalamus neuron	Stromal cell	Astrocytes_Aldoc^{high}	Neuron_cellcytle^{high}	Oligodendrocyte precursors	Radial glia cell_Neurog2^{high}	Radial glia cell_Gdf10 high	Astrocytes_Ccnb1^{high}	Radial glia cell_Id3^{high}	Neuron_Fam19a2^{high}	Endothelial cell	Cajal-Retzius neuron cells	GABAergic neuron_Lhx6^{high}	GABAergic neuron_Htra3^{high}
Macrophage/Microglia	100	0	0	0	0	0	0	0	0	0	5	1	0	0	0	0	0	1	0	5	0	0	0
Ependymal cells	0	100	1	31	9	6	0	14	1	1	0	0	0	0	28	0	0	0	0	0	4	0	1
GABAergic neuron_Nrxn3^{high}	0	1	100	2	4	2	5	0	2	3	0	0	0	0	0	0	0	0	1	0	1	54	29
Cortical neurons	0	31	2	100	28	33	1	6	24	5	0	0	0	0	2	0	0	0	3	0	2	1	2
GABA-Interneurons	0	9	4	28	100	24	10	3	29	11	0	0	0	0	1	0	0	0	5	0	1	6	6
Erythrocyte	0	6	2	33	24	100	4	3	46	5	0	0	0	0	0	0	0	0	4	0	0	0	2
Dopaminergic amacrine neuron	0	0	5	1	10	4	100	1	17	17	0	0	0	1	0	0	0	0	10	0	6	12	7
Immature neuron	0	14	0	6	3	3	1	100	0	13	1	0	0	2	11	4	0	0	0	0	5	0	5
Glutamatergic neuron	0	1	2	24	29	46	17	0	100	8	0	0	0	0	0	0	0	1	11	0	1	3	4
Hypothalamus neuron	0	1	3	5	11	5	17	13	8	100	0	0	0	0	0	1	0	1	2	0	2	5	3
Stromal cell	5	0	0	0	0	0	0	1	0	0	100	2	0	3	0	2	1	5	0	20	2	0	0
Astrocytes_Aldoc^{high}	1	0	0	0	0	0	0	0	0	0	2	100	0	9	14	25	21	39	0	2	0	0	1
Neuron_cellcytle^{high}	0	0	0	0	0	0	0	0	0	0	0	0	100	3	2	0	63	2	0	0	0	0	0
Oligodendrocyte precursors	0	0	0	0	0	0	1	2	0	0	3	9	3	100	7	6	3	16	0	1	2	1	1
Radial glia cell_Neurog2^{high}	0	28	0	2	1	0	0	11	0	0	0	14	2	7	100	12	2	4	0	0	0	0	3
Radial glia cell_Gdf10^{high}	0	0	0	0	0	0	0	4	0	1	2	25	0	6	12	100	9	30	0	1	3	0	0
Astrocytes_Ccnb1^{high}	0	0	0	0	0	0	0	0	0	0	1	21	63	3	2	9	100	12	0	0	0	0	0
Radial glia cell_Id3^{high}	1	0	0	0	0	0	0	0	1	1	5	39	2	16	4	30	12	100	0	2	1	0	1
Neuron_Fam19a2^{high}	0	0	1	3	5	4	10	0	11	2	0	0	0	0	0	0	0	0	100	0	0	2	3
Endothelial cell	5	0	0	0	0	0	0	0	0	0	20	2	0	1	0	1	0	2	0	100	1	0	0
Cajal-Retzius neuron cells	0	4	1	2	1	0	6	5	1	2	2	0	0	2	0	3	0	1	0	1	100	2	2
GABAergic neuron_Lhx6^{high}	0	0	54	1	6	0	12	0	3	5	0	0	0	1	0	0	0	0	2	0	2	100	33
GABAergic neuron_Htra3^{high}	0	1	29	2	6	2	7	5	4	3	0	1	0	1	3	0	0	1	3	0	2	33	100

References

Abrahams, B. S., Arking, D. E., Campbell, D. B., Mefford, H. C., Morrow, E. M., Weiss, L. A., Menashe, I., Wadkins, T., Banerjee-Basu, S. and Packer, A. 2013. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism,* 4**,** 36.

Agirman, G., Broix, L. and Nguyen, L. 2017. Cerebral cortex development: an outside-in perspective. *FEBS Lett,* 591**,** 3978-3992.

Akbarian, S., Liu, C., Knowles, J. A., Vaccarino, F. M., Farnham, P. J., Crawford, G. E., Jaffe, A. E., Pinto, D., Dracheva, S., Geschwind, D. H., et al. 2015. The PsychENCODE project. *Nat Neurosci,* 18**,** 1707-12.

Alanis-Lobato, G., Andrade-Navarro, M. A. and Schaefer, M. H. 2017. HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res,* 45**,** D408-D414.

Alver, R. C., Chadha, G. S., Gillespie, P. J. and Blow, J. J. 2017. Reversal of DDK-Mediated MCM Phosphorylation by Rif1-PP1 Regulates Replication Initiation and Replisome Stability Independently of ATR/Chk1. *Cell Rep,* 18**,** 2508-2520.

Angevine, J. B., Jr. and Sidman, R. L. 1961. Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse. *Nature,* 192**,** 766-8.

Aran, D., Hu, Z. and Butte, A. J. 2017. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol,* 18**,** 220.

Atladottir, H. O., Thorsen, P., Ostergaard, L., Schendel, D. E., Lemcke, S., Abdallah, M. and Parner, E. T. 2010. Maternal infection requiring hospitalization during pregnancy and autism spectrum disorders. *J Autism Dev Disord,* 40**,** 1423-30.

Attalla, S., Taifour, T., Bui, T. and Muller, W. 2020. Insights from transgenic mouse models of PyMT-induced breast cancer: recapitulating human breast cancer progression in vivo. *Oncogene.*

Aulmann, S., Adler, N., Rom, J., Helmchen, B., Schirmacher, P. and Sinn, H. P. 2006. c-myc amplifications in primary breast carcinomas and their local recurrences. *J Clin Pathol,* 59**,** 424-8.

Awate, S. and De Benedetti, A. 2016. TLK1B mediated phosphorylation of Rad9 regulates its nuclear/cytoplasmic localization and cell cycle checkpoint. *BMC Mol Biol,* 17**,** 3.

Bakail, M., Gaubert, A., Andreani, J., Moal, G., Pinna, G., Boyarchuk, E., Gaillard, M. C., Courbeyrette, R., Mann, C., Thuret, J. Y., et al. 2019. Design on a Rational Basis of High-Affinity Peptides Inhibiting the Histone Chaperone ASF1. *Cell Chem Biol,* 26**,** 1573-1585 e10.

Bandeira, F., Lent, R. and Herculano-Houzel, S. 2009. Changing numbers of neuronal and non-neuronal cells underlie postnatal brain growth in the rat. *Proc Natl Acad Sci U S A,* 106**,** 14108-13.

Banerjee-Basu, S. and Packer, A. 2010. SFARI Gene: an evolving database for the autism research community. *Dis Model Mech,* 3**,** 133-5.

Bankhead, P., Loughrey, M. B., Fernandez, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., et al. 2017. QuPath: Open source software for digital pathology image analysis. *Sci Rep,* 7**,** 16878.

Bao, W., Kojima, K. K. and Kohany, O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob DNA,* 6**,** 11.

Barbar, E. 2008. Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. *Biochemistry,* 47**,** 503-8.

Barlow, C., Hirotsune, S., Paylor, R., Liyanage, M., Eckhaus, M., Collins, F., Shiloh, Y., Crawley, J. N., Ried, T., Tagle, D., et al. 1996. Atm-deficient mice: a paradigm of ataxia telangiectasia. *Cell,* 86**,** 159-71.

Barnes, T. A. and Amir, E. 2017. HYPE or HOPE: the prognostic value of infiltrating immune cells in cancer. *Br J Cancer,* 117**,** 451-460.

Bartkova, J., Horejsi, Z., Koed, K., Kramer, A., Tort, F., Zieger, K., Guldberg, P., Sehested, M., Nesland, J. M., Lukas, C., et al. 2005. DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature,* 434**,** 864-70.

Batsukh, T., Schulz, Y., Wolf, S., Rabe, T. I., Oellerich, T., Urlaub, H., Schaefer, I. M. and Pauli, S. 2012. Identification and characterization of FAM124B as a novel component of a CHD7 and CHD8 containing complex. *PLoS One,* 7**,** e52640.

Becker, J. R., Cuella-Martin, R., Barazas, M., Liu, R., Oliveira, C., Oliver, A. W., Bilham, K., Holt, A. B., Blackford, A. N., Heierhorst, J., et al. 2018. The ASCIZ-DYNLL1 axis promotes 53BP1-dependent non-homologous end joining and PARP inhibitor sensitivity. *Nat Commun,* 9**,** 5406.

Benjamini Y., a. H. Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society,* 57**,** 289-300.

Bennardo, N., Cheng, A., Huang, N. and Stark, J. M. 2008. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet,* 4**,** e1000110.

Berg, J. M. and Geschwind, D. H. 2012. Autism genetics: searching for specificity and convergence. *Genome Biol,* 13**,** 247.

Bernier, R., Golzio, C., Xiong, B., Stessman, H. A., Coe, B. P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A. T., et al. 2014. Disruptive CHD8 mutations define a subtype of autism early in development. *Cell,* 158**,** 263-276.

Bertoli, C., Skotheim, J. M. and de Bruin, R. A. 2013. Control of cell cycle transcription during G1 and S phases. *Nat Rev Mol Cell Biol,* 14**,** 518-28.

Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., Coussens, L. M., Gabrilovich, D. I., Ostrand-Rosenberg, S., Hedrick, C. C., et al. 2018. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med,* 24**,** 541-550.

Blackford, A. N. and Jackson, S. P. 2017. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell,* 66**,** 801-817.

Blasius, M., Forment, J. V., Thakkar, N., Wagner, S. A., Choudhary, C. and Jackson, S. P. 2011. A phospho-proteomic screen identifies substrates of the checkpoint kinase Chk1. *Genome Biol,* 12**,** R78.

Boldt, K., van Reeuwijk, J., Lu, Q., Koutroumpas, K., Nguyen, T. M., Texier, Y., van Beersum, S. E., Horn, N., Willer, J. R., Mans, D. A., et al. 2016. An organelle-specific protein landscape identifies novel diseases and molecular mechanisms. *Nat Commun,* 7**,** 11491.

Bouwman, P. and Jonkers, J. 2012. The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance. *Nat Rev Cancer,* 12**,** 587-98.

Box, G. E. P., Jenkins, G.M., and Reinsel G.C. 1994. Time Series Analysis: Forecasting and Control. *Holden-Day,* 3rd Edition.

Brahmer, J. R., Tykodi, S. S., Chow, L. Q., Hwu, W. J., Topalian, S. L., Hwu, P., Drake, C. G., Camacho, L. H., Kauh, J., Odunsi, K., et al. 2012. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med,* 366**,** 2455-65.

Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N. and Ting, A. Y. 2018. Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol,* 36**,** 880-887.

Brown, A. S. 2006. Prenatal infection as a risk factor for schizophrenia. *Schizophr Bull,* 32**,** 200-2.

Brown, R. E., Stanford, L. and Schellinck, H. M. 2000. Developing standardized behavioral tests for knockout and mutant mice. *ILAR J,* 41**,** 163-74.

Bruni, D., Angell, H. K. and Galon, J. 2020. The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. *Nat Rev Cancer,* 20**,** 662-680.

Bryant, H. E., Petermann, E., Schultz, N., Jemth, A. S., Loseva, O., Issaeva, N., Johansson, F., Fernandez, S., McGlynn, P. and Helleday, T. 2009. PARP is activated at stalled forks to mediate Mre11-dependent replication restart and recombination. *EMBO J,* 28**,** 2601-15.

Bryant, L., Li, D., Cox, S. G., Marchione, D., Joiner, E. F., Wilson, K., Janssen, K., Lee, P., March, M. E., Nair, D., et al. 2020. Histone H3.3 beyond cancer: Germline mutations in Histone 3 Family 3A and 3B cause a previously unidentified neurodegenerative disorder in 46 patients. *Sci Adv,* 6.

Burstyn, I., Wang, X., Yasui, Y., Sithole, F. and Zwaigenbaum, L. 2011. Autism spectrum disorders and fetal hypoxia in a population-based cohort: accounting for missing exposures via Estimation-Maximization algorithm. *BMC Med Res Methodol,* 11**,** 2.

Byun, T. S., Pacek, M., Yee, M. C., Walter, J. C. and Cimprich, K. A. 2005. Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev,* 19**,** 1040-52.

Caballé Mestres, A., Berenguer Llergo, A., and Stephan-Otto Attolini, C. 2018. Adjusting for systematic technical biases in risk assessment of gene signatures in transcriptomic cancer cohorts. *bioRxiv*.

Cai, Z., Chehab, N. H. and Pavletich, N. P. 2009. Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase. *Mol Cell,* 35**,** 818-29.

Canfield, C., Rains, J. and De Benedetti, A. 2009. TLK1B promotes repair of DSBs via its interaction with Rad9 and Asf1. *BMC Mol Biol,* 10**,** 110.

Cardoso, A. R., Lopes-Marques, M., Silva, R. M., Serrano, C., Amorim, A., Prata, M. J. and Azevedo, L. 2019. Essential genetic findings in neurodevelopmental disorders. *Hum Genomics,* 13**,** 31.

Carlson, M. 2018. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.7.0.

Carpentier, P. A., Dingman, A. L. and Palmer, T. D. 2011. Placental TNF-alpha signaling in illness-induced complications of pregnancy. *Am J Pathol,* 178**,** 2802-10.

Carrera, P., Moshkin, Y. M., Gronke, S., Sillje, H. H., Nigg, E. A., Jackle, H. and Karch, F. 2003. Tousled-like kinase functions with the chromatin assembly pathway regulating nuclear divisions. *Genes Dev,* 17**,** 2578-90.

Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N. and Szallasi, Z. 2006. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet,* 38**,** 1043-8.

Carvalho, B. S. and Irizarry, R. A. 2010. A framework for oligonucleotide microarray preprocessing. *Bioinformatics,* 26**,** 2363-7.

Chanas-Sacre, G., Rogister, B., Moonen, G. and Leprince, P. 2000. Radial glia phenotype: origin, regulation, and transdifferentiation. *J Neurosci Res,* 61**,** 357-63.

Chapman, J. R., Barral, P., Vannier, J. B., Borel, V., Steger, M., Tomas-Loba, A., Sartori, A. A., Adams, I. R., Batista, F. D. and Boulton, S. J. 2013. RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol Cell,* 49**,** 858-71.

Chen, V. S., Morrison, J. P., Southwell, M. F., Foley, J. F., Bolon, B. and Elmore, S. A. 2017. Histology Atlas of the Developing Prenatal and Postnatal Mouse Central Nervous System, with Emphasis on Prenatal Days E7.5 to E18.5. *Toxicol Pathol,* 45**,** 705-744.

Cherian, M. P. 2004. Seckel-like syndrome or Seckel variants? *Ann Saudi Med,* 24**,** 469-72.

Chess, S. 1971. Autism in children with congenital rubella. *J Autism Child Schizophr,* 1**,** 33-47.

Cheung, H. H., Pei, D. and Chan, W. Y. 2015. Stem cell aging in adult progeria. *Cell Regen,* 4**,** 6.

Choi, G. B., Yim, Y. S., Wong, H., Kim, S., Kim, H., Kim, S. V., Hoeffer, C. A., Littman, D. R. and Huh, J. R. 2016. The maternal interleukin-17a pathway in mice promotes autism-like phenotypes in offspring. *Science,* 351**,** 933-9.

Christensen, J., Gronborg, T. K., Sorensen, M. J., Schendel, D., Parner, E. T., Pedersen, L. H. and Vestergaard, M. 2013. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA,* 309**,** 1696-703.

Ciccia, A. and Elledge, S. J. 2010. The DNA damage response: making it safe to play with knives. *Mol Cell,* 40**,** 179-204.

Cimprich, K. A. and Cortez, D. 2008. ATR: an essential regulator of genome integrity. *Nat Rev Mol Cell Biol,* 9**,** 616-27.

Cohen, P. and Alessi, D. R. 2013. Kinase drug discovery--what's next in the field? *ACS Chem Biol,* 8**,** 96-104.

Coquel, F., Silva, M. J., Techer, H., Zadorozhny, K., Sharma, S., Nieminuszczy, J., Mettling, C., Dardillac, E., Barthe, A., Schmitz, A. L., et al. 2018. SAMHD1 acts at stalled replication forks to prevent interferon induction. *Nature,* 557**,** 57-61.

Cornacchia, D., Dileep, V., Quivy, J. P., Foti, R., Tili, F., Santarella-Mellwig, R., Antony, C., Almouzni, G., Gilbert, D. M. and Buonomo, S. B. 2012. Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J,* 31**,** 3678-90.

Corominas, R., Yang, X., Lin, G. N., Kang, S., Shen, Y., Ghamsari, L., Broly, M., Rodriguez, M., Tam, S., Trigg, S. A., et al. 2014. Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nat Commun,* 5**,** 3650.

Corpet, A., De Koning, L., Toedling, J., Savignoni, A., Berger, F., Lemaitre, C., O'Sullivan, R. J., Karlseder, J., Barillot, E., Asselain, B., et al. 2011. Asf1b, the necessary Asf1 isoform for proliferation, is predictive of outcome in breast cancer. *EMBO J,* 30**,** 480-93.

Cotney, J., Muhle, R. A., Sanders, S. J., Liu, L., Willsey, A. J., Niu, W., Liu, W., Klei, L., Lei, J., Yin, J., et al. 2015. The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat Commun,* 6**,** 6404.

Cristino, A. S., Williams, S. M., Hawi, Z., An, J. Y., Bellgrove, M. A., Schwartz, C. E., Costa Lda, F. and Claudianos, C. 2014. Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. *Mol Psychiatry,* 19**,** 294-301.

Crotti, A. and Ransohoff, R. M. 2016. Microglial Physiology and Pathophysiology: Insights from Genome-wide Transcriptional Profiling. *Immunity,* 44**,** 505-515.

Cunningham, C. L., Martinez-Cerdeno, V. and Noctor, S. C. 2013. Microglia regulate the number of neural precursor cells in the developing cerebral cortex. *J Neurosci,* 33**,** 4216-33.

Danussi, C., Bose, P., Parthasarathy, P. T., Silberman, P. C., Van Arnam, J. S., Vitucci, M., Tang, O. Y., Heguy, A., Wang, Y., Chan, T. A., et al. 2018. Atrx inactivation drives disease-defining phenotypes in glioma cells of origin through global epigenomic remodeling. *Nat Commun,* 9**,** 1057.

Davie, S. A., Maglione, J. E., Manner, C. K., Young, D., Cardiff, R. D., MacLeod, C. L. and Ellies, L. G. 2007. Effects of FVB/NJ and C57Bl/6J strain backgrounds on

mammary tumor phenotype in inducible nitric oxide synthase deficient mice. *Transgenic Res,* 16**,** 193-201.

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature,* 515**,** 209-15.

Delacroix, S., Wagner, J. M., Kobayashi, M., Yamamoto, K. and Karnitz, L. M. 2007. The Rad9-Hus1-Rad1 (9-1-1) clamp activates checkpoint signaling via TopBP1. *Genes Dev,* 21**,** 1472-7.

DeNardo, D. G., Barreto, J. B., Andreu, P., Vasquez, L., Tawfik, D., Kolhatkar, N. and Coussens, L. M. 2009. CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell,* 16**,** 91-102.

Dheen, S. T., Kaur, C. and Ling, E. A. 2007. Microglial activation and its implications in the brain diseases. *Curr Med Chem,* 14**,** 1189-97.

Di Wu, E. L., François Vaillant, Marie-Liesse Asselin-Labat, Jane E. Visvader, Gordon K. Smyth 2010. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics,* 26**,** 2176-2182.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

Driscoll, D. J. O., Felice, V. D., Kenny, L. C., Boylan, G. B. and O'Keeffe, G. W. 2018. Mild prenatal hypoxia-ischemia leads to social deficits and central and peripheral inflammation in exposed offspring. *Brain Behav Immun,* 69**,** 418-427.

Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., Lydon, N. B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., et al. 2001. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med,* 344**,** 1031-7.

Du, Z., Podsypanina, K., Huang, S., McGrath, A., Toneff, M. J., Bogoslovskaia, E., Zhang, X., Moraes, R. C., Fluck, M., Allred, D. C., et al. 2006. Introduction of oncogenes into mammary glands in vivo with an avian retroviral vector initiates and promotes carcinogenesis in mouse models. *Proc Natl Acad Sci U S A,* 103**,** 17396-401.

Duan, Q., Zhang, H., Zheng, J. and Zhang, L. 2020. Turning Cold into Hot: Firing up the Tumor Microenvironment. *Trends Cancer,* 6**,** 605-618.

Duchon, A. and Herault, Y. 2016. DYRK1A, a Dosage-Sensitive Gene Involved in Neurodevelopmental Disorders, Is a Target for Drug Development in Down Syndrome. *Front Behav Neurosci,* 10**,** 104.

Earl, R. K., Turner, T. N., Mefford, H. C., Hudac, C. M., Gerdts, J., Eichler, E. E. and Bernier, R. A. 2017. Clinical phenotype of ASD-associated DYRK1A haploinsufficiency. *Mol Autism,* 8**,** 54.

Efron B, R. 2007. On testing the significance of sets of genes. *Annals of Applied Statistics,* 1**,** 107-129.

Eklund, A. C. and Szallasi, Z. 2008. Correction of technical bias in clinical microarray data improves concordance with known biological information. *Genome Biol,* 9**,** R26.

Elsasser, S. J., Noh, K. M., Diaz, N., Allis, C. D. and Banaszynski, L. A. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature,* 522**,** 240-244.

Erblich, B., Zhu, L., Etgen, A. M., Dobrenis, K. and Pollard, J. W. 2011. Absence of colony stimulation factor-1 receptor results in loss of microglia, disrupted brain development and olfactory deficits. *PLoS One,* 6**,** e26317.

Escribano-Diaz, C., Orthwein, A., Fradet-Turcotte, A., Xing, M., Young, J. T., Tkac, J., Cook, M. A., Rosebrock, A. P., Munro, M., Canny, M. D., et al. 2013. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell,* 49**,** 872-83.

Estes, M. L. and McAllister, A. K. 2016. Maternal immune activation: Implications for neuropsychiatric disorders. *Science,* 353**,** 772-7.

Fan, G., Martinowich, K., Chin, M. H., He, F., Fouse, S. D., Hutnick, L., Hattori, D., Ge, W., Shen, Y., Wu, H., et al. 2005. DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development,* 132**,** 3345-56.

Felker, A., Nieuwenhuize, S., Dolbois, A., Blazkova, K., Hess, C., Low, L. W., Burger, S., Samson, N., Carney, T. J., Bartunek, P., et al. 2016. In Vivo Performance and Properties of Tamoxifen Metabolites for CreERT2 Control. *PLoS One,* 11**,** e0152989.

Felter BC, d. F. P. J., Miyamoto KN, Bonatto D 2016. Genome Instability: from virus to human application. *Academic Press***,** 447-462.

Feng, J., Zhou, Y., Campbell, S. L., Le, T., Li, E., Sweatt, J. D., Silva, A. J. and Fan, G. 2010. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nat Neurosci,* 13**,** 423-30.

Feng, W., Kawauchi, D., Korkel-Qu, H., Deng, H., Serger, E., Sieber, L., Lieberman, J. A., Jimeno-Gonzalez, S., Lambo, S., Hanna, B. S., et al. 2017a. Chd7 is indispensable for mammalian brain development through activation of a neuronal differentiation programme. *Nat Commun,* 8**,** 14758.

Feng, W., Shao, C. and Liu, H. K. 2017b. Versatile Roles of the Chromatin Remodeler CHD7 during Brain Development and Disease. *Front Mol Neurosci,* 10**,** 309.

Ferguson, F. M. and Gray, N. S. 2018. Kinase inhibitors: the road ahead. *Nat Rev Drug Discov,* 17**,** 353-377.

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R. M. and Carter, N. P. 2009. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet,* 84**,** 524-33.

Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M. and Kolodner, R. 1993. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell,* 75**,** 1027-38.

Fisher, R. A. 1922. n the interpretation of χ2 from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society,* 85**,** 87-94.

Flex, E., Martinelli, S., Van Dijck, A., Ciolfi, A., Cecchetti, S., Coluzzi, E., Pannone, L., Andreoli, C., Radio, F. C., Pizzi, S., et al. 2019. Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. *Am J Hum Genet,* 105**,** 493-508.

Fluck, M. M. and Schaffhausen, B. S. 2009. Lessons in signaling and tumorigenesis from polyomavirus middle T antigen. *Microbiol Mol Biol Rev,* 73**,** 542-63, Table of Contents.

Forment, J. V. and O'Connor, M. J. 2018. Targeting the replication stress response in cancer. *Pharmacol Ther,* 188**,** 155-167.

Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein, F. A., Huber, W., Johnstone, E., et al. 2016. Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program. *Mol Cell,* 61**,** 260-73.

Gajewski, T. F. 2015. The Next Hurdle in Cancer Immunotherapy: Overcoming the Non-T-Cell-Inflamed Tumor Microenvironment. *Semin Oncol,* 42**,** 663-71.

Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., Schork, A. J., Appadurai, V., Buil, A., Werge, T. M., et al. 2018a. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science,* 359**,** 693-697.

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., Won, H., van Bakel, H., Varghese, M., Wang, Y., et al. 2018b. Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science,* 362.

Ganesan, H., Balasubramanian, V., Iyer, M., Venugopal, A., Subramaniam, M. D., Cho, S. G. and Vellingiri, B. 2019. mTOR signalling pathway - A root cause for idiopathic autism? *BMB Rep,* 52**,** 424-433.

Gao, Y., Li, P. and Pappas, D. 2013. A microfluidic localized, multiple cell culture array using vacuum actuated cell seeding: integrated anticancer drug testing. *Biomed Microdevices,* 15**,** 907-15.

Garrote, A. M., Redondo, P., Montoya, G. and Munoz, I. G. 2014. Purification, crystallization and preliminary X-ray diffraction analysis of the kinase domain of human tousled-like kinase 2. *Acta Crystallogr F Struct Biol Commun,* 70**,** 354-7.

Ge, W. P., Miyawaki, A., Gage, F. H., Jan, Y. N. and Jan, L. Y. 2012. Local generation of glia is a major astrocyte source in postnatal cortex. *Nature,* 484**,** 376-80.

Gentleman, R., Carey V., Huber, W., Irizarry, R., Dudoit, S 2005. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, Springer, New York.

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol,* 5**,** R80.

Ginhoux, F., Greter, M., Leboeuf, M., Nandi, S., See, P., Gokhan, S., Mehler, M. F., Conway, S. J., Ng, L. G., Stanley, E. R., et al. 2010. Fate mapping analysis reveals that adult microglia derive from primitive macrophages. *Science,* 330**,** 841-5.

Glass, C. K., Saijo, K., Winner, B., Marchetto, M. C. and Gage, F. H. 2010. Mechanisms underlying inflammation in neurodegeneration. *Cell,* 140**,** 918-34.

Gonatopoulos-Pournatzis, T., Wu, M., Braunschweig, U., Roth, J., Han, H., Best, A. J., Raj, B., Aregger, M., O'Hanlon, D., Ellis, J. D., et al. 2018. Genome-wide CRISPR-Cas9 Interrogation of Splicing Networks Reveals a Mechanism for Recognition of Autism-Misregulated Neuronal Microexons. *Mol Cell,* 72**,** 510-524 e12.

Gonzalez-Roca, E., Garcia-Albeniz, X., Rodriguez-Mulero, S., Gomis, R. R., Kornacker, K. and Auer, H. 2010. Accurate expression profiling of very small cell populations. *PLoS One,* 5**,** e14418.

Gorgoulis, V. G., Vassiliou, L. V., Karakaidos, P., Zacharatos, P., Kotsinas, A., Liloglou, T., Venere, M., Ditullio, R. A., Jr., Kastrinakis, N. G., Levy, B., et al. 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature,* 434**,** 907-13.

Gotz, M. and Barde, Y. A. 2005. Radial glial cells defined and major intermediates between embryonic stem cells and CNS neurons. *Neuron,* 46**,** 369-72.

Gotz, M. and Huttner, W. B. 2005. The cell biology of neurogenesis. *Nat Rev Mol Cell Biol,* 6**,** 777-88.

Grossman, R. L., Heath, A. P., Ferretti, V., Varmus, H. E., Lowy, D. R., Kibbe, W. A. and Staudt, L. M. 2016. Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med,* 375**,** 1109-12.

Groth, A., Lukas, J., Nigg, E. A., Sillje, H. H., Wernstedt, C., Bartek, J. and Hansen, K. 2003. Human Tousled like kinases are targeted by an ATM- and Chk1-dependent DNA damage checkpoint. *EMBO J,* 22**,** 1676-87.

Groth, A., Ray-Gallet, D., Quivy, J. P., Lukas, J., Bartek, J. and Almouzni, G. 2005. Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol Cell,* 17**,** 301-11.

Gruber, M., Hu, C. J., Johnson, R. S., Brown, E. J., Keith, B. and Simon, M. C. 2007. Acute postnatal ablation of Hif-2alpha results in anemia. *Proc Natl Acad Sci U S A,* 104**,** 2301-6.

Günesdogan, U., Jackle, H. and Herzig, A. 2014. Histone supply regulates S phase timing and cell cycle progression. *Elife,* 3**,** e02443.

Gupta, G. P., Vanness, K., Barlas, A., Manova-Todorova, K. O., Wen, Y. H. and Petrini, J. H. 2013. The Mre11 complex suppresses oncogene-driven breast tumorigenesis and metastasis. *Mol Cell,* 52**,** 353-65.

Gupta, R., Somyajit, K., Narita, T., Maskey, E., Stanlie, A., Kremer, M., Typas, D., Lammers, M., Mailand, N., Nussenzweig, A., et al. 2018. DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell,* 173**,** 972-988 e23.

Gupta, S., Ellis, S. E., Ashar, F. N., Moes, A., Bader, J. S., Zhan, J., West, A. B. and Arking, D. E. 2014. Transcriptome analysis reveals dysregulation of innate immune response genes and neuronal activity-dependent genes in autism. *Nat Commun,* 5**,** 5748.

Guy, C. T., Cardiff, R. D. and Muller, W. J. 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol,* 12**,** 954-61.

Halazonetis, T. D., Gorgoulis, V. G. and Bartek, J. 2008. An oncogene-induced DNA damage model for cancer development. *Science,* 319**,** 1352-5.

Hamdan, F. F., Daoud, H., Rochefort, D., Piton, A., Gauthier, J., Langlois, M., Foomani, G., Dobrzeniecka, S., Krebs, M. O., Joober, R., et al. 2010. De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet,* 87**,** 671-8.

Hamdan, F. F., Srour, M., Capo-Chichi, J. M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O., et al. 2014. De novo mutations in moderate or severe intellectual disability. *PLoS Genet,* 10**,** e1004772.

Hammond, C. M., Stromme, C. B., Huang, H., Patel, D. J. and Groth, A. 2017. Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol,* 18**,** 141-158.

Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye, F., et al. 2018. Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell,* 173**,** 1307.

Han, Z., Riefler, G. M., Saam, J. R., Mango, S. E. and Schumacher, J. M. 2005. The C. elegans Tousled-like kinase contributes to chromosome segregation as a substrate and regulator of the Aurora B kinase. *Curr Biol,* 15**,** 894-904.

Han, Z., Saam, J. R., Adams, H. P., Mango, S. E. and Schumacher, J. M. 2003. The C. elegans Tousled-like kinase (TLK-1) has an essential role in transcription. *Curr Biol,* 13**,** 1921-9.

Hanahan, D. and Weinberg, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

Havrilla, J. M., Pedersen, B. S., Layer, R. M. and Quinlan, A. R. 2019. A map of constrained coding regions in the human genome. *Nat Genet,* 51**,** 88-95.

Hayes, D. F. 2019. HER2 and Breast Cancer - A Phenomenal Success Story. *N Engl J Med,* 381**,** 1284-1286.

He, F., Ge, W., Martinowich, K., Becker-Catania, S., Coskun, V., Zhu, W., Wu, H., Castro, D., Guillemot, F., Fan, G., et al. 2005. A positive autoregulatory loop of Jak-STAT signaling controls the onset of astrogliogenesis. *Nat Neurosci,* 8**,** 616-25.

He, Y. J., Meghani, K., Caron, M. C., Yang, C., Ronato, D. A., Bian, J., Sharma, A., Moore, J., Niraj, J., Detappe, A., et al. 2018. DYNLL1 binds to MRE11 to limit DNA end resection in BRCA1-deficient cells. *Nature,* 563**,** 522-526.

Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., et al. 2015. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell,* 163**,** 712-23.

Heyn, P., Logan, C. V., Fluteau, A., Challis, R. C., Auchynnikava, T., Martin, C. A., Marsh, J. A., Taglini, F., Kilanowski, F., Parry, D. A., et al. 2019. Gain-of-function DNMT3A mutations cause microcephalic dwarfism and hypermethylation of Polycomb-regulated regions. *Nat Genet,* 51**,** 96-105.

Hiraga, S., Alvino, G. M., Chang, F., Lian, H. Y., Sridhar, A., Kubota, T., Brewer, B. J., Weinreich, M., Raghuraman, M. K. and Donaldson, A. D. 2014. Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex. *Genes Dev,* 28**,** 372-83.

Hoang, S. M. and O'Sullivan, R. J. 2020. Alternative Lengthening of Telomeres: Building Bridges To Connect Chromosome Ends. *Trends Cancer,* 6**,** 247-260.

Hothorn, T., Bretz, F. and Westfall, P. 2008. Simultaneous inference in general parametric models. *Biom J,* 50**,** 346-63.

Hrdlickova, R., Toloue, M. and Tian, B. 2017. RNA-Seq methods for transcriptome analysis. *Wiley Interdiscip Rev RNA,* 8.

Hsiao, E. Y. and Patterson, P. H. 2011. Activation of the maternal immune system induces endocrine changes in the placenta via IL-6. *Brain Behav Immun,* 25**,** 604-15.

Huertas, P. 2010. DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol,* 17**,** 11-6.

Hurst, T. P. and Magiorkinis, G. 2015. Activation of the innate immune response by endogenous retroviruses. *J Gen Virol,* 96**,** 1207-1218.

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., Stessman, H. A., Witherspoon, K. T., Vives, L., Patterson, K. E., et al. 2014. The contribution of de novo coding mutations to autism spectrum disorder. *Nature,* 515**,** 216-21.

Ira, G. and Nussenzweig, A. 2014. A new Riff: Rif1 eats its cake and has it too. *EMBO Rep,* 15**,** 622-4.

Irimia, M., Weatheritt, R. J., Ellis, J. D., Parikshak, N. N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallieres, M., Tapial, J., Raj, B., O'Hanlon, D., et al. 2014. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell,* 159**,** 1511-23.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. and Speed, T. P. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics,* 4**,** 249-64.

Ito, D., Imai, Y., Ohsawa, K., Nakajima, K., Fukuuchi, Y. and Kohsaka, S. 1998. Microglia-specific localisation of a novel calcium binding protein, Iba1. *Brain Res Mol Brain Res,* 57**,** 1-9.

Jackson, S. P. and Bartek, J. 2009. The DNA-damage response in human biology and disease. *Nature,* 461**,** 1071-8.

Jang, C. W., Shibata, Y., Starmer, J., Yee, D. and Magnuson, T. 2015. Histone H3.3 maintains genome integrity during mammalian development. *Genes Dev,* 29**,** 1377-92.

Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G. C., Lukas, J. and Jackson, S. P. 2006. ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol,* 8**,** 37-45.

Jeanne, M., Vuillaume, M. L., Ung, D. C., Vancollie, V. E., Wagner, C., Collins, S. C., Vonwill, S., Haye, D., Chelloug, N., Pfundt, R., et al. 2021. Haploinsufficiency of the HIRA gene located in the 22q11 deletion syndrome region is associated with abnormal neurodevelopment and impaired dendritic outgrowth. *Hum Genet*.

Jonsson, M. E., Garza, R., Johansson, P. A. and Jakobsson, J. 2020. Transposable Elements: A Common Feature of Neurodevelopmental and Neurodegenerative Disorders. *Trends Genet,* 36**,** 610-623.

Jonsson, M. E., Garza, R., Sharma, Y., Petri, R., Sodersten, E., Johansson, J. G., Johansson, P. A., Atacho, D. A., Pircs, K., Madsen, S., et al. 2021. Activation of endogenous retroviruses during brain development causes an inflammatory response. *EMBO J***,** e106423.

Kalay, E., Yigit, G., Aslan, Y., Brown, K. E., Pohl, E., Bicknell, L. S., Kayserili, H., Li, Y., Tuysuz, B., Nurnberg, G., et al. 2011. CEP152 is a genome maintenance protein disrupted in Seckel syndrome. *Nat Genet,* 43**,** 23-6.

Kallioniemi, A., Kallioniemi, O. P., Piper, J., Tanner, M., Stokke, T., Chen, L., Smith, H. S., Pinkel, D., Gray, J. W. and Waldman, F. M. 1994. Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci U S A,* 91**,** 2156-60.

Kanehisa, M. and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res,* 28**,** 27-30.

Kang, P., Lee, H. K., Glasgow, S. M., Finley, M., Donti, T., Gaber, Z. B., Graham, B. H., Foster, A. E., Novitch, B. G., Gronostajski, R. M., et al. 2012. Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis. *Neuron,* 74**,** 79-94.

Kataoka, S., Takuma, K., Hara, Y., Maeda, Y., Ago, Y. and Matsuda, T. 2013. Autism-like behaviours with transient histone hyperacetylation in mice treated prenatally with valproic acid. *Int J Neuropsychopharmacol,* 16**,** 91-103.

Kelemen, L. E., Wang, X., Fredericksen, Z. S., Pankratz, V. S., Pharoah, P. D., Ahmed, S., Dunning, A. M., Easton, D. F., Vierkant, R. A., Cerhan, J. R., et al. 2009. Genetic variation in the chromosome 17q23 amplicon and breast cancer risk. *Cancer Epidemiol Biomarkers Prev,* 18**,** 1864-8.

Kelly, R. and Davey, S. K. 2013. Tousled-like kinase-dependent phosphorylation of Rad9 plays a role in cell cycle progression and G2/M checkpoint exit. *PLoS One,* 8**,** e85859.

Kim, J. A., Tan, Y., Wang, X., Cao, X., Veeraraghavan, J., Liang, Y., Edwards, D. P., Huang, S., Pan, X., Li, K., et al. 2016. Comprehensive functional analysis of the tousled-like kinase 2 frequently amplified in aggressive luminal breast cancers. *Nat Commun,* 7**,** 12991.

Kim, S. 2015. ppcor: An R package for a fast calculation to semi-partial cor- relation coefficients. *Commun. Stat. Appl. Methods*, 665–674.

Kim, S., Kim, H., Yim, Y. S., Ha, S., Atarashi, K., Tan, T. G., Longman, R. S., Honda, K., Littman, D. R., Choi, G. B., et al. 2017. Maternal gut bacteria promote neurodevelopmental abnormalities in mouse offspring. *Nature,* 549**,** 528-532.

King, C. R., Kraus, M. H. and Aaronson, S. A. 1985. Amplification of a novel v-erbB-related gene in a human mammary carcinoma. *Science,* 229**,** 974-6.

Klimovskaia, I. M., Young, C., Stromme, C. B., Menard, P., Jasencakova, Z., Mejlvang, J., Ask, K., Ploug, M., Nielsen, M. L., Jensen, O. N., et al. 2014. Tousled-like kinases phosphorylate Asf1 to promote histone supply during DNA replication. *Nat Commun,* 5**,** 3394.

Knight, J. D. R., Choi, H., Gupta, G. D., Pelletier, L., Raught, B., Nesvizhskii, A. I. and Gingras, A. C. 2017. ProHits-viz: a suite of web tools for visualizing interaction proteomics data. *Nat Methods,* 14**,** 645-646.

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R. and Massouras, A. 2019. VarSome: the human genomic variant search engine. *Bioinformatics,* 35**,** 1978-1980.

Kotsantis, P., Petermann, E. and Boulton, S. J. 2018. Mechanisms of Oncogene-Induced Replication Stress: Jigsaw Falling into Place. *Cancer Discov,* 8**,** 537-555.

Krakowiak, P., Goines, P. E., Tancredi, D. J., Ashwood, P., Hansen, R. L., Hertz-Picciotto, I. and Van de Water, J. 2017. Neonatal Cytokine Profiles Associated With Autism Spectrum Disorder. *Biol Psychiatry,* 81**,** 442-451.

Krause, D. R., Jonnalagadda, J. C., Gatei, M. H., Sillje, H. H., Zhou, B. B., Nigg, E. A. and Khanna, K. 2003. Suppression of Tousled-like kinase activity after DNA damage or replication block requires ATM, NBS1 and Chk1. *Oncogene,* 22**,** 5927-37.

Kuhn, S., Gritti, L., Crooks, D. and Dombrowski, Y. 2019. Oligodendrocytes in Development, Myelin Generation and Beyond. *Cells,* 8.

Lalani, S. R., Safiullah, A. M., Fernbach, S. D., Harutyunyan, K. G., Thaller, C., Peterson, L. E., McPherson, J. D., Gibbs, R. A., White, L. D., Hefner, M., et al. 2006. Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and genotype-phenotype correlation. *Am J Hum Genet,* 78**,** 303-14.

Lavery, L. A., Ure, K., Wan, Y. W., Luo, C., Trostle, A. J., Wang, W., Jin, H., Lopez, J., Lucero, J., Durham, M. A., et al. 2020. Losing Dnmt3a dependent methylation in inhibitory neurons impairs neural function by a mechanism impacting Rett syndrome. *Elife,* 9.

Le Naour, J., Zitvogel, L., Galluzzi, L., Vacchelli, E. and Kroemer, G. 2020. Trial watch: STING agonists in cancer therapy. *Oncoimmunology,* 9**,** 1777624.

Lecona, E. and Fernandez-Capetillo, O. 2018. Targeting ATR in cancer. *Nat Rev Cancer,* 18**,** 586-595.

Lee, K. Y., Im, J. S., Shibata, E. and Dutta, A. 2017. ASF1a Promotes Non-homologous End Joining Repair by Facilitating Phosphorylation of MDC1 by ATM at Double-Strand Breaks. *Mol Cell,* 68**,** 61-75 e5.

Lee, S. B., Segura-Bayona, S., Villamor-Paya, M., Saredi, G., Todd, M. A. M., Attolini, C. S., Chang, T. Y., Stracker, T. H. and Groth, A. 2018. Tousled-like kinases stabilize replication forks and show synthetic lethality with checkpoint and PARP inhibitors. *Sci Adv,* 4**,** eaat4985.

Lelieveld, S. H., Reijnders, M. R., Pfundt, R., Yntema, H. G., Kamsteeg, E. J., de Vries, P., de Vries, B. B., Willemsen, M. H., Kleefstra, T., Lohner, K., et al. 2016. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci,* 19**,** 1194-6.

Lengauer, C., Kinzler, K. W. and Vogelstein, B. 1998. Genetic instabilities in human cancers. *Nature,* 396**,** 643-9.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y. H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K., et al. 2011. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron,* 70**,** 886-97.

Li, B. and Dewey, C. N. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics,* 12**,** 323.

Li, Y., DeFatta, R., Anthony, C., Sunavala, G. and De Benedetti, A. 2001. A translationally regulated Tousled kinase phosphorylates histone H3 and confers radioresistance when overexpressed. *Oncogene,* 20**,** 726-38.

Li, Y. and Jiao, J. 2017. Histone chaperone HIRA regulates neural progenitor cell proliferation and neurogenesis via beta-catenin. *J Cell Biol,* 216**,** 1975-1992.

Li, Z., Gourguechon, S. and Wang, C. C. 2007. Tousled-like kinase in a microbial eukaryote regulates spindle assembly and S-phase progression by interacting with Aurora kinase and chromatin assembly factors. *J Cell Sci,* 120**,** 3883-94.

Liao, X., Liu, Y., Fu, X. and Li, Y. 2020. Postmortem Studies of Neuroinflammation in Autism Spectrum Disorder: a Systematic Review. *Mol Neurobiol,* 57**,** 3424-3438.

Liao, Y., Smyth, G. K. and Shi, W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res,* 41**,** e108.

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P. and Tamayo, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* 1**,** 417-425.

Lim, S. and Kaldis, P. 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development,* 140**,** 3079-93.

Lin, E. Y., Jones, J. G., Li, P., Zhu, L., Whitney, K. D., Muller, W. J. and Pollard, J. W. 2003. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol,* 163**,** 2113-26.

Lin, E. Y., Nguyen, A. V., Russell, R. G. and Pollard, J. W. 2001. Colony-stimulating factor 1 promotes progression of mammary tumors to malignancy. *J Exp Med,* 193**,** 727-40.

Lin, M., Yao, Z., Zhao, N. and Zhang, C. 2019. TLK2 enhances aggressive phenotypes of glioblastoma cells through the activation of SRC signaling pathway. *Cancer Biol Ther,* 20**,** 101-108.

Linnerbauer, M., Wheeler, M. A. and Quintana, F. J. 2020. Astrocyte Crosstalk in CNS Inflammation. *Neuron,* 108**,** 608-622.

Liu, Q., Yin, X., Languino, L. R. and Altieri, D. C. 2018. Evaluation of drug combination effect using a Bliss independence dose-response surface model. *Stat Biopharm Res,* 10**,** 112-122.

Lo, K. W., Kan, H. M., Chan, L. N., Xu, W. G., Wang, K. P., Wu, Z., Sheng, M. and Zhang, M. 2005. The 8-kDa dynein light chain binds to p53-binding protein 1 and mediates DNA damage-induced p53 nuclear accumulation. *J Biol Chem,* 280**,** 8172-9.

Love, M. I., Huber, W. and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

Luskin, M. B. and Shatz, C. J. 1985. Studies of the earliest generated cells of the cat's visual cortex: cogeneration of subplate and marginal zones. *J Neurosci,* 5**,** 1062-75.

Lynch, H. T., Lynch, P. M., Lanspa, S. J., Snyder, C. L., Lynch, J. F. and Boland, C. R. 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet,* 76**,** 1-18.

Macheret, M. and Halazonetis, T. D. 2015. DNA replication stress as a hallmark of cancer. *Annu Rev Pathol,* 10**,** 425-48.

Mackenzie, K. J., Carroll, P., Martin, C. A., Murina, O., Fluteau, A., Simpson, D. J., Olova, N., Sutcliffe, H., Rainger, J. K., Leitch, A., et al. 2017. cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature,* 548**,** 461-465.

Majewski, F. and Goecke, T. 1982. Studies of microcephalic primordial dwarfism I: approach to a delineation of the Seckel syndrome. *Am J Med Genet,* 12**,** 7-21.

Malatesta, P., Hartfuss, E. and Gotz, M. 2000. Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage. *Development,* 127**,** 5253-63.

Malhotra, D., McCarthy, S., Michaelson, J. J., Vacic, V., Burdick, K. E., Yoon, S., Cichon, S., Corvin, A., Gary, S., Gershon, E. S., et al. 2011. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron,* 72**,** 951-63.

Malkova, N. V., Yu, C. Z., Hsiao, E. Y., Moore, M. J. and Patterson, P. H. 2012. Maternal immune activation yields offspring displaying mouse versions of the three core symptoms of autism. *Brain Behav Immun,* 26**,** 607-16.

Malumbres, M. 2014. Cyclin-dependent kinases. *Genome Biol,* 15**,** 122.

Manning, B. J. and Yusufzai, T. 2017. The ATP-dependent chromatin remodeling enzymes CHD6, CHD7, and CHD8 exhibit distinct nucleosome binding and remodeling activities. *J Biol Chem,* 292**,** 11927-11936.

Marie, C., Clavairoly, A., Frah, M., Hmidan, H., Yan, J., Zhao, C., Van Steenwinckel, J., Daveau, R., Zalc, B., Hassan, B., et al. 2018. Oligodendrocyte precursor survival and differentiation requires chromatin remodeling by Chd7 and Chd8. *Proc Natl Acad Sci U S A,* 115**,** E8246-E8255.

Masi, A., Quintana, D. S., Glozier, N., Lloyd, A. R., Hickie, I. B. and Guastella, A. J. 2015. Cytokine aberrations in autism spectrum disorder: a systematic review and meta-analysis. *Mol Psychiatry,* 20**,** 440-6.

Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., 3rd, Hurov, K. E., Luo, J., Bakalarski, C. E., Zhao, Z., Solimini, N., Lerenthal, Y., et al. 2007. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science,* 316**,** 1160-6.

Maver, A., Cuturilo, G., Ruml, S. J. and Peterlin, B. 2019. Clinical Next Generation Sequencing Reveals an H3F3A Gene as a New Potential Gene Candidate for Microcephaly Associated with Severe Developmental Delay, Intellectual Disability and Growth Retardation. *Balkan J Med Genet,* 22**,** 65-68.

Maze, I., Wenderski, W., Noh, K. M., Bagot, R. C., Tzavaras, N., Purushothaman, I., Elsasser, S. J., Guo, Y., Ionete, C., Hurd, Y. L., et al. 2015. Critical Role of Histone Turnover in Neuronal Transcription and Plasticity. *Neuron,* 87**,** 77-94.

McConnell, S. K. 1995. Constructing the cerebral cortex: neurogenesis and fate determination. *Neuron,* 15**,** 761-8.

McKinnon, P. J. 2012. ATM and the molecular pathogenesis of ataxia telangiectasia. *Annu Rev Pathol,* 7**,** 303-21.

Mertins, P., Mani, D. R., Ruggles, K. V., Gillette, M. A., Clauser, K. R., Wang, P., Wang, X., Qiao, J. W., Cao, S., Petralia, F., et al. 2016. Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature,* 534**,** 55-62.

Mi, H., Muruganujan, A., Ebert, D., Huang, X. and Thomas, P. D. 2019. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res,* 47**,** D419-D426.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science,* 266**,** 66-71.

Miknis, G. F., Stevens, S. J., Smith, L. E., Ostrov, D. A. and Churchill, M. E. 2015. Development of novel Asf1-H3/H4 inhibitors. *Bioorg Med Chem Lett,* 25**,** 963-8.

Min, W., Bruhn, C., Grigaravicius, P., Zhou, Z. W., Li, F., Kruger, A., Siddeek, B., Greulich, K. O., Popp, O., Meisezahl, C., et al. 2013. Poly(ADP-ribose) binding to Chk1 at stalled replication forks is required for S-phase checkpoint activation. *Nat Commun,* 4**,** 2993.

Misson, J. P., Austin, C. P., Takahashi, T., Cepko, C. L. and Caviness, V. S., Jr. 1991. The alignment of migrating neural cells in relation to the murine neopallial radial glial fiber system. *Cereb Cortex,* 1**,** 221-9.

Miyata, T. 2008. Development of three-dimensional architecture of the neuroepithelium: role of pseudostratification and cellular 'community'. *Dev Growth Differ,* 50 Suppl 1**,** S105-12.

Moll, U. M. and Petrenko, O. 2003. The MDM2-p53 interaction. *Mol Cancer Res,* 1**,** 1001-8.

Molyneaux, B. J., Arlotta, P., Menezes, J. R. and Macklis, J. D. 2007. Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci,* 8**,** 427-37.

Moncini, S., Bedeschi, M. F., Castronovo, P., Crippa, M., Calvello, M., Garghentino, R. R., Scuvera, G., Finelli, P. and Venturin, M. 2013. ATRX mutation in two adult brothers with non-specific moderate intellectual disability identified by exome sequencing. *Meta Gene,* 1**,** 102-8.

Morgan, J. T., Chana, G., Pardo, C. A., Achim, C., Semendeferi, K., Buckwalter, J., Courchesne, E. and Everall, I. P. 2010. Microglial activation and increased microglial density observed in the dorsolateral prefrontal cortex in autism. *Biol Psychiatry,* 68**,** 368-76.

Mortuza, G. B., Hermida, D., Pedersen, A. K., Segura-Bayona, S., Lopez-Mendez, B., Redondo, P., Ruther, P., Pozdnyakova, I., Garrote, A. M., Munoz, I. G., et al. 2018. Molecular basis of Tousled-Like Kinase 2 activation. *Nat Commun,* 9**,** 2535.

Moynahan, M. E., Chiu, J. W., Koller, B. H. and Jasin, M. 1999. Brca1 controls homology-directed DNA repair. *Mol Cell,* 4**,** 511-8.

Moynahan, M. E., Pierce, A. J. and Jasin, M. 2001. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell,* 7**,** 263-72.

Murga, M., Bunting, S., Montana, M. F., Soria, R., Mulero, F., Canamero, M., Lee, Y., McKinnon, P. J., Nussenzweig, A. and Fernandez-Capetillo, O. 2009. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat Genet,* 41**,** 891-8.

Murga, M., Campaner, S., Lopez-Contreras, A. J., Toledo, L. I., Soria, R., Montana, M. F., Artista, L., Schleker, T., Guerra, C., Garcia, E., et al. 2011. Exploiting oncogene-induced replicative stress for the selective killing of Myc-driven tumors. *Nat Struct Mol Biol,* 18**,** 1331-1335.

Nam, E. A. and Cortez, D. 2011. ATR signalling: more than meeting at the fork. *Biochem J,* 436**,** 527-36.

Namihira, M., Kohyama, J., Semi, K., Sanosaka, T., Deneen, B., Taga, T. and Nakashima, K. 2009. Committed neuronal precursors confer astrocytic potential on residual neural precursor cells. *Dev Cell,* 16**,** 245-55.

Nandi, S., Gokhan, S., Dai, X. M., Wei, S., Enikolopov, G., Lin, H., Mehler, M. F. and Stanley, E. R. 2012. The CSF-1 receptor ligands IL-34 and CSF-1 exhibit distinct developmental brain expression patterns and regulate neural progenitor cell maintenance and maturation. *Dev Biol,* 367**,** 100-13.

Napoli, I. and Neumann, H. 2009. Microglial clearance function in health and disease. *Neuroscience,* 158**,** 1030-8.

Napolioni, V., Ober-Reynolds, B., Szelinger, S., Corneveaux, J. J., Pawlowski, T., Ober-Reynolds, S., Kirwan, J., Persico, A. M., Melmed, R. D., Craig, D. W., et al. 2013. Plasma cytokine profiling in sibling pairs discordant for autism spectrum disorder. *J Neuroinflammation,* 10**,** 38.

Negrini, S., Gorgoulis, V. G. and Halazonetis, T. D. 2010. Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol,* 11**,** 220-8.

Niemi, M. E. K., Martin, H. C., Rice, D. L., Gallone, G., Gordon, S., Kelemen, M., McAloney, K., McRae, J., Radford, E. J., Yu, S., et al. 2018. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature,* 562**,** 268-271.

Nikodemova, M., Kimyon, R. S., De, I., Small, A. L., Collier, L. S. and Watters, J. J. 2015. Microglial numbers attain adult levels after undergoing a rapid decrease in cell number in the third postnatal week. *J Neuroimmunol,* 278**,** 280-8.

Noctor, S. C., Flint, A. C., Weissman, T. A., Dammerman, R. S. and Kriegstein, A. R. 2001. Neurons derived from radial glial cells establish radial units in neocortex. *Nature,* 409**,** 714-20.

Noctor, S. C., Martinez-Cerdeno, V., Ivic, L. and Kriegstein, A. R. 2004. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat Neurosci,* 7**,** 136-44.

O'Roak, B. J., Deriziotis, P., Lee, C., Vives, L., Schwartz, J. J., Girirajan, S., Karakoc, E., Mackenzie, A. P., Ng, S. B., Baker, C., et al. 2011. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet,* 43**,** 585-9.

Oh, T., Fakurnejad, S., Sayegh, E. T., Clark, A. J., Ivan, M. E., Sun, M. Z., Safaee, M., Bloch, O., James, C. D. and Parsa, A. T. 2014. Immunocompetent murine models for the study of glioblastoma immunotherapy. *J Transl Med,* 12**,** 107.

Ohtsuka, T. and Kageyama, R. 2019. Regulation of temporal properties of neural stem cells and transition timing of neurogenesis and gliogenesis during mammalian neocortical development. *Semin Cell Dev Biol,* 95**,** 4-11.

Oughtred, R., Stark, C., Breitkreutz, B. J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. 2019. The BioGRID interaction database: 2019 update. *Nucleic Acids Res,* 47**,** D529-D541.

Pacek, M. and Walter, J. C. 2004. A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *EMBO J,* 23**,** 3667-76.

Pages, F., Galon, J., Dieu-Nosjean, M. C., Tartour, E., Sautes-Fridman, C. and Fridman, W. H. 2010. Immune infiltration in human tumors: a prognostic factor that should not be ignored. *Oncogene,* 29**,** 1093-102.

Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D., et al. 1994. Mutation of a mutL homolog in hereditary colon cancer. *Science,* 263**,** 1625-9.

Pardo, C. A., Vargas, D. L. and Zimmerman, A. W. 2005. Immunity, neuroglia and neuroinflammation in autism. *Int Rev Psychiatry,* 17**,** 485-95.

Parenti, I., Rabaneda, L. G., Schoen, H. and Novarino, G. 2020. Neurodevelopmental Disorders: From Genetics to Functional Pathways. *Trends Neurosci,* 43**,** 608-621.

Parikshak, N. N., Luo, R., Zhang, A., Won, H., Lowe, J. K., Chandran, V., Horvath, S. and Geschwind, D. H. 2013. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell,* 155**,** 1008-21.

Patterson, P. H. 2011. Maternal infection and immune involvement in autism. *Trends Mol Med,* 17**,** 389-94.

Pavinato, L., Villamor-Paya, M., Sanchiz-Calvo, M., Andreoli, C., Gay, M., Vilaseca, M., Arauz-Garofalo, G., Ciolfi, A., Bruselles, A., Pippucci, T., et al. 2020. Functional analysis of TLK2 variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis. *J Med Genet*.

Perez, C., Sawmiller, D. and Tan, J. 2016. The role of heparan sulfate deficiency in autistic phenotype: potential involvement of Slit/Robo/srGAPs-mediated dendritic spine formation. *Neural Dev,* 11**,** 11.

Petitprez, F., Meylan, M., de Reynies, A., Sautes-Fridman, C. and Fridman, W. H. 2020. The Tumor Microenvironment in the Response to Immune Checkpoint Blockade Therapies. *Front Immunol,* 11**,** 784.

Pinheiro, J. C., and Bates, D.M. 2000. Mixed-Effects Models in S and S-PLUS. *Springer,* esp**,** 236, 243.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., et al. 2014. Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *Am J Hum Genet,* 94**,** 677-94.

Premoli, M., Bonini, S. A., Mastinu, A., Maccarinelli, G., Aria, F., Paiardi, G. and Memo, M. 2019. Specific profile of ultrasonic communication in a mouse model of neurodevelopmental disorders. *Sci Rep,* 9**,** 15912.

Quesnel-Vallieres, M., Dargaei, Z., Irimia, M., Gonatopoulos-Pournatzis, T., Ip, J. Y., Wu, M., Sterne-Weiler, T., Nakagawa, S., Woodin, M. A., Blencowe, B. J., et al. 2016. Misregulation of an Activity-Dependent Splicing Network as a Common Mechanism Underlying Autism Spectrum Disorders. *Mol Cell,* 64**,** 1023-1034.

Quesnel-Vallieres, M., Irimia, M., Cordes, S. P. and Blencowe, B. J. 2015. Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. *Genes Dev,* 29**,** 746-59.

Qvist, P., Huertas, P., Jimeno, S., Nyegaard, M., Hassan, M. J., Jackson, S. P. and Borglum, A. D. 2011. CtIP Mutations Cause Seckel and Jawad Syndromes. *PLoS Genet,* 7**,** e1002310.

Rapali, P., Szenes, A., Radnai, L., Bakos, A., Pal, G. and Nyitray, L. 2011. DYNLL/LC8: a light chain subunit of the dynein motor complex and beyond. *FEBS J,* 278**,** 2980-96.

Regue, L., Sdelci, S., Bertran, M. T., Caelles, C., Reverter, D. and Roig, J. 2011. DYNLL/LC8 protein controls signal transduction through the Nek9/Nek6 signaling module by regulating Nek6 binding to Nek9. *J Biol Chem,* 286**,** 18118-29.

Reid, D. A., Reed, P. J., Schlachetzki, J. C. M., Nitulescu, II, Chou, G., Tsui, E. C., Jones, J. R., Chandran, S., Lu, A. T., McClain, C. A., et al. 2021. Incorporation of a nucleoside analog maps genome repair sites in postmitotic human neurons. *Science,* 372**,** 91-94.

Reijnders, M. R. F., Miller, K. A., Alvi, M., Goos, J. A. C., Lees, M. M., de Burca, A., Henderson, A., Kraus, A., Mikat, B., de Vries, B. B. A., et al. 2018. De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder. *Am J Hum Genet,* 102**,** 1195-1203.

Rein, K., Yanez, D. A., Terre, B., Palenzuela, L., Aivio, S., Wei, K., Edelmann, W., Stark, J. M. and Stracker, T. H. 2015. EXO1 is critical for embryogenesis and the DNA damage response in mice with a hypomorphic Nbs1 allele. *Nucleic Acids Res,* 43**,** 7371-87.

Richet, N., Liu, D., Legrand, P., Velours, C., Corpet, A., Gaubert, A., Bakail, M., Moal-Raisin, G., Guerois, R., Compper, C., et al. 2015. Structural insight into how the human helicase subunit MCM2 may act as a histone chaperone together with ASF1 at the replication fork. *Nucleic Acids Res,* 43**,** 1905-17.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* 43**,** e47.

Roe, J. L., Durfee, T., Zupan, J. R., Repetti, P. P., McLean, B. G. and Zambryski, P. C. 1997. TOUSLED is a nuclear serine/threonine protein kinase that requires a coiled-coil region for oligomerization and catalytic activity. *J Biol Chem,* 272**,** 5838-45.

Roe, J. L., Rivin, C. J., Sessions, R. A., Feldmann, K. A. and Zambryski, P. C. 1993. The Tousled gene in A. thaliana encodes a protein kinase homolog that is required for leaf and flower development. *Cell,* 75**,** 939-50.

Ronald, S., Awate, S., Rath, A., Carroll, J., Galiano, F., Dwyer, D., Kleiner-Hancock, H., Mathis, J. M., Vigod, S. and De Benedetti, A. 2013. Phenothiazine Inhibitors of TLKs Affect Double-Strand Break Repair and DNA Damage Response Recovery and Potentiate Tumor Killing with Radiomimetic Therapy. *Genes Cancer,* 4**,** 39-53.

Roskoski, R., Jr. 2020. Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacol Res,* 152**,** 104609.

Rossell, D. 2009. GaGa: a parsimonious and flexible model for differential expression analysis. . *Annals of Applied Statistics,* 3**,** 1035-1051.

Rossell, D. 2013. gaga: GaGa hierarchical model for high-throughput data analysis. R package version 2.10.0.

Rossell, D., Stephan-Otto Attolini, C., Kroiss, M. and Stocker, A. 2014. Quantifying Alternative Splicing from Paired-End Rna-Sequencing Data. *Ann Appl Stat,* 8**,** 309-330.

Rothblum-Oviatt, C., Wright, J., Lefton-Greif, M. A., McGrath-Morrow, S. A., Crawford, T. O. and Lederman, H. M. 2016. Ataxia telangiectasia: a review. *Orphanet J Rare Dis,* 11**,** 159.

Roux, K. J., Kim, D. I., Raida, M. and Burke, B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol,* 196**,** 801-10.

Sadic, D., Schmidt, K., Groh, S., Kondofersky, I., Ellwart, J., Fuchs, C., Theis, F. J. and Schotta, G. 2015. Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep,* 16**,** 836-50.

Sakai, Y., Shaw, C. A., Dawson, B. C., Dugas, D. V., Al-Mohtaseb, Z., Hill, D. E. and Zoghbi, H. Y. 2011. Protein interactome reveals converging molecular pathways among autism disorders. *Sci Transl Med,* 3**,** 86ra49.

Saleh, A., Macia, A. and Muotri, A. R. 2019. Transposable Elements, Inflammation, and Neurological Disease. *Front Neurol,* 10**,** 894.

Sanders, S. J., Murtha, M. T., Gupta, A. R., Murdoch, J. D., Raubeson, M. J., Willsey, A. J., Ercan-Sencicek, A. G., DiLullo, N. M., Parikshak, N. N., Stein, J. L., et al. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature,* 485**,** 237-41.

Sato, A. 2016. mTOR, a Potential Target to Treat Autism Spectrum Disorder. *CNS Neurol Disord Drug Targets,* 15**,** 533-43.

Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., Peng, M., Collins, R., Grove, J., Klei, L., et al. 2020. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell,* 180**,** 568-584 e23.

Schafer, D. P. and Stevens, B. 2013. Phagocytic glial cells: sculpting synaptic circuits in the developing nervous system. *Curr Opin Neurobiol,* 23**,** 1034-40.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods,* 9**,** 676-82.

Schluth-Bolard, C., Diguet, F., Chatron, N., Rollat-Farnier, P. A., Bardel, C., Afenjar, A., Amblard, F., Amiel, J., Blesson, S., Callier, P., et al. 2019. Whole genome paired-end sequencing elucidates functional and phenotypic consequences of balanced chromosomal rearrangement in patients with developmental disorders. *J Med Genet,* 56**,** 526-535.

Schmechel, D. E. and Rakic, P. 1979. A Golgi study of radial glial cells in developing monkey telencephalon: morphogenesis and transformation into astrocytes. *Anat Embryol (Berl),* 156**,** 115-52.

Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., et al. 2007. Strong association of de novo copy number mutations with autism. *Science,* 316**,** 445-9.

Segura-Bayona, S. 2018. Role of the Tousled like kinases in maintaining genome and epigenome stability. *Doctoral thesis.*

Segura-Bayona, S., Knobel, P. A., Gonzalez-Buron, H., Youssef, S. A., Pena-Blanco, A., Coyaud, E., Lopez-Rovira, T., Rein, K., Palenzuela, L., Colombelli, J., et al. 2017.

Differential requirements for Tousled-like kinases 1 and 2 in mammalian development. *Cell Death Differ,* 24**,** 1872-1885.

Segura-Bayona, S. and Stracker, T. H. 2019. The Tousled-like kinases regulate genome and epigenome stability: implications in development and disease. *Cell Mol Life Sci,* 76**,** 3827-3841.

Segura-Bayona, S., Villamor-Paya, M., Attolini, C. S., Koenig, L. M., Sanchiz-Calvo, M., Boulton, S. J. and Stracker, T. H. 2020. Tousled-Like Kinases Suppress Innate Immune Signaling Triggered by Alternative Lengthening of Telomeres. *Cell Rep,* 32**,** 107983.

Senturk, E. and Manfredi, J. J. 2013. p53 and cell cycle effects after DNA damage. *Methods Mol Biol,* 962**,** 49-61.

Seol, J. H., Song, T. Y., Oh, S. E., Jo, C., Choi, A., Kim, B., Park, J., Hong, S., Song, I., Jung, K. Y., et al. 2015. Identification of small molecules that inhibit the histone chaperone Asf1 and its chromatin function. *BMB Rep,* 48**,** 685-90.

Sessa, A., Mao, C. A., Hadjantonakis, A. K., Klein, W. H. and Broccoli, V. 2008. Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex. *Neuron,* 60**,** 56-69.

Shaltiel, I. A., Krenning, L., Bruinsma, W. and Medema, R. H. 2015. The same, only different - DNA damage checkpoints and their reversal throughout the cell cycle. *J Cell Sci,* 128**,** 607-20.

Sharma, P. and Allison, J. P. 2015. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. *Cell,* 161**,** 205-14.

Sharpe, A. H. and Pauken, K. E. 2018. The diverse functions of the PD1 inhibitory pathway. *Nat Rev Immunol,* 18**,** 153-167.

Shi, T., Bunker, R. D., Mattarocci, S., Ribeyre, C., Faty, M., Gut, H., Scrima, A., Rass, U., Rubin, S. M., Shore, D., et al. 2013. Rif1 and Rif2 shape telomere function and architecture through multivalent Rap1 interactions. *Cell,* 153**,** 1340-53.

Silljé, H. H. and Nigg, E. A. 2001. Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases. *Curr Biol,* 11**,** 1068-73.

Silljé, H. H., Takahashi, K., Tanaka, K., Van Houwe, G. and Nigg, E. A. 1999. Mammalian homologues of the plant Tousled gene code for cell-cycle-regulated kinases with maximal activities linked to ongoing DNA replication. *EMBO J,* 18**,** 5691-702.

Silverman, J. L., Yang, M., Lord, C. and Crawley, J. N. 2010. Behavioural phenotyping assays for mouse models of autism. *Nat Rev Neurosci,* 11**,** 490-502.

Sinclair, C. S., Rowley, M., Naderi, A. and Couch, F. J. 2003. The 17q23 amplicon and breast cancer. *Breast Cancer Res Treat,* 78**,** 313-22.

Singh, T., Walters, J. T. R., Johnstone, M., Curtis, D., Suvisaari, J., Torniainen, M., Rees, E., Iyegbe, C., Blackwood, D., McIntosh, A. M., et al. 2017a. The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nat Genet,* 49**,** 1167-1173.

Singh, V., Bhoir, S., Chikhale, R. V., Hussain, J., Dwyer, D., Bryce, R. A., Kirubakaran, S. and De Benedetti, A. 2020. Generation of Phenothiazine with Potent Anti-TLK1 Activity for Prostate Cancer Therapy. *iScience,* 23**,** 101474.

Singh, V., Connelly, Z. M., Shen, X. and De Benedetti, A. 2017b. Identification of the proteome complement of humanTLK1 reveals it binds and phosphorylates NEK1 regulating its activity. *Cell Cycle,* 16**,** 915-926.

Singh, V., Jaiswal, P. K., Ghosh, I., Koul, H. K., Yu, X. and De Benedetti, A. 2019a. Targeting the TLK1/NEK1 DDR axis with Thioridazine suppresses outgrowth of androgen independent prostate tumors. *Int J Cancer,* 145**,** 1055-1067.

Singh, V., Jaiswal, P. K., Ghosh, I., Koul, H. K., Yu, X. and De Benedetti, A. 2019b. The TLK1-Nek1 axis promotes prostate cancer progression. *Cancer Lett,* 453**,** 131-141.

Sir, J. H., Barr, A. R., Nicholas, A. K., Carvalho, O. P., Khurshid, M., Sossick, A., Reichelt, S., D'Santos, C., Woods, C. G. and Gergely, F. 2011. A primary microcephaly protein complex forms a ring around parental centrioles. *Nat Genet,* 43**,** 1147-53.

Sirbu, B. M., Couch, F. B., Feigerle, J. T., Bhaskara, S., Hiebert, S. W. and Cortez, D. 2011. Analysis of protein dynamics at active, stalled, and collapsed replication forks. *Genes Dev,* 25**,** 1320-7.

Smit, A., Hubley, R & Green, P. 2013-2015. RepeatMasker Open-4.0. <http://www.repeatmasker.org>.

Solek, C. M., Farooqi, N., Verly, M., Lim, T. K. and Ruthazer, E. S. 2018. Maternal immune activation in neurodevelopmental disorders. *Dev Dyn,* 247**,** 588-619.

Sollis, E., Graham, S. A., Vino, A., Froehlich, H., Vreeburg, M., Dimitropoulou, D., Gilissen, C., Pfundt, R., Rappold, G. A., Brunner, H. G., et al. 2016. Identification and functional characterization of de novo FOXP1 variants provides novel insights into the etiology of neurodevelopmental disorder. *Hum Mol Genet,* 25**,** 546-57.

Solovyov, A., Vabret, N., Arora, K. S., Snyder, A., Funt, S. A., Bajorin, D. F., Rosenberg, J. E., Bhardwaj, N., Ting, D. T. and Greenbaum, B. D. 2018. Global Cancer Transcriptome Quantifies Repeat Element Polarization between Immunotherapy Responsive and T Cell Suppressive Classes. *Cell Rep,* 23**,** 512-521.

Srivastava, M., Chen, Z., Zhang, H., Tang, M., Wang, C., Jung, S. Y. and Chen, J. 2018. Replisome Dynamics and Their Functional Relevance upon DNA Damage through the PCNA Interactome. *Cell Rep,* 25**,** 3869-3883 e4.

Stracker, T. H. and Petrini, J. H. 2011. The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol,* 12**,** 90-103.

Stracker, T. H., Roig, I., Knobel, P. A. and Marjanovic, M. 2013. The ATM signaling network in development and disease. *Front Genet,* 4**,** 37.

Straughen, J. K., Misra, D. P., Divine, G., Shah, R., Perez, G., VanHorn, S., Onbreyt, V., Dygulska, B., Schmitt, R., Lederman, S., et al. 2017. The association between placental histopathology and autism spectrum disorder. *Placenta,* 57**,** 183-188.

Suhasini, A. N. and Brosh, R. M., Jr. 2013. Disease-causing missense mutations in human DNA helicase disorders. *Mutat Res,* 752**,** 138-152.

Sukackaite, R., Cornacchia, D., Jensen, M. R., Mas, P. J., Blackledge, M., Enervald, E., Duan, G., Auchynnikava, T., Kohn, M., Hart, D. J., et al. 2017. Mouse Rif1 is a regulatory subunit of protein phosphatase 1 (PP1). *Sci Rep,* 7**,** 2119.

Sunavala-Dossabhoy, G. 2018. Preserving salivary gland physiology against genotoxic damage - the Tousled way. *Oral Dis,* 24**,** 1390-1398.

Sunavala-Dossabhoy, G. and De Benedetti, A. 2009. Tousled homolog, TLK1, binds and phosphorylates Rad9; TLK1 acts as a molecular chaperone in DNA repair. *DNA Repair (Amst),* 8**,** 87-102.

Sunavala-Dossabhoy, G., Fowler, M. and De Benedetti, A. 2004. Translation of the radioresistance kinase TLK1B is induced by gamma-irradiation through activation of mTOR and phosphorylation of 4E-BP1. *BMC Mol Biol,* 5**,** 1.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., et al. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res,* 47**,** D607-D613.

Takata, A., Miyake, N., Tsurusaki, Y., Fukai, R., Miyatake, S., Koshimizu, E., Kushima, I., Okada, T., Morikawa, M., Uno, Y., et al. 2018. Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. *Cell Rep,* 22**,** 734-747.

Tapial, J., Ha, K. C. H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallieres, M., Permanyer, J., Sodaei, R., Marquez, Y., et al. 2017. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res,* 27**,** 1759-1768.

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. and Prins, P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics,* 31**,** 2032-4.

Taylor, A. M., Shih, J., Ha, G., Gao, G. F., Zhang, X., Berger, A. C., Schumacher, S. E., Wang, C., Hu, H., Liu, J., et al. 2018. Genomic and Functional Approaches to Understanding Cancer Aneuploidy. *Cancer Cell,* 33**,** 676-689 e3.

Team, R. C. 2017. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A. C. and Choi, H. 2014. SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics,* 100**,** 37-43.

Terabayashi, T. and Hanada, K. 2018. Genome instability syndromes caused by impaired DNA repair and aberrant DNA damage responses. *Cell Biol Toxicol,* 34**,** 337-350.

Terré, B., Lewis, M., Gil-Gomez, G., Han, Z., Lu, H., Aguilera, M., Prats, N., Roy, S., Zhao, H. and Stracker, T. H. 2019. Defects in efferent duct multiciliogenesis underlie male infertility in GEMC1-, MCIDAS- or CCNO-deficient mice. *Development,* 146.

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., Porta-Pardo, E., Gao, G. F., Plaisier, C. L., Eddy, J. A., et al. 2018. The Immune Landscape of Cancer. *Immunity,* 48**,** 812-830 e14.

Toledo, L. I., Murga, M., Zur, R., Soria, R., Rodriguez, A., Martinez, S., Oyarzabal, J., Pastor, J., Bischoff, J. R. and Fernandez-Capetillo, O. 2011. A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nat Struct Mol Biol,* 18**,** 721-7.

Tong, C. K. and Vidyadaran, S. 2016. Role of microglia in embryonic neurogenesis. *Exp Biol Med (Maywood),* 241**,** 1669-75.

Topalian, S. L., Hodi, F. S., Brahmer, J. R., Gettinger, S. N., Smith, D. C., McDermott, D. F., Powderly, J. D., Carvajal, R. D., Sosman, J. A., Atkins, M. B., et al. 2012. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med,* 366**,** 2443-54.

Topf, A., Oktay, Y., Balaraju, S., Yilmaz, E., Sonmezler, E., Yis, U., Laurie, S., Thompson, R., Roos, A., MacArthur, D. G., et al. 2020. Severe neurodevelopmental disease caused by a homozygous TLK2 variant. *Eur J Hum Genet,* 28**,** 383-387.

Torné, J., Ray-Gallet, D., Boyarchuk, E., Garnier, M., Le Baccon, P., Coulon, A., Orsi, G. A. and Almouzni, G. 2020. Two HIRA-dependent pathways mediate H3.3 de novo deposition and recycling during transcription. *Nat Struct Mol Biol,* 27**,** 1057-1068.

Treiber, D. K. and Shah, N. P. 2013. Ins and outs of kinase DFG motifs. *Chem Biol,* 20**,** 745-6.

Tremont, A., Lu, J. and Cole, J. T. 2017. Endocrine Therapy for Early Breast Cancer: Updated Review. *Ochsner J,* 17**,** 405-411.

Tronnes, A. A., Koschnitzky, J., Daza, R., Hitti, J., Ramirez, J. M. and Hevner, R. 2016. Effects of Lipopolysaccharide and Progesterone Exposures on Embryonic Cerebral Cortex Development in Mice. *Reprod Sci,* 23**,** 771-8.

Udugama, M., FT, M. C., Chan, F. L., Tang, M. C., Pickett, H. A., JD, R. M., Mayne, L., Collas, P., Mann, J. R. and Wong, L. H. 2015. Histone variant H3.3 provides the heterochromatic H3 lysine 9 tri-methylation mark at telomeres. *Nucleic Acids Res,* 43**,** 10227-37.

van Zundert, G. C. P., Rodrigues, J., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J. and Bonvin, A. 2016. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol,* 428**,** 720-725.

Vargas, D. L., Nascimbene, C., Krishnan, C., Zimmerman, A. W. and Pardo, C. A. 2005. Neuroglial activation and neuroinflammation in the brain of patients with autism. *Ann Neurol,* 57**,** 67-81.

Vijayan, R. S., He, P., Modi, V., Duong-Ly, K. C., Ma, H., Peterson, J. R., Dunbrack, R. L., Jr. and Levy, R. M. 2015. Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. *J Med Chem,* 58**,** 466-79.

Voigt, T. 1989. Development of glial cells in the cerebral wall of ferrets: direct tracing of their transformation from radial glia into astrocytes. *J Comp Neurol,* 289**,** 74-88.

Voineagu, I., Wang, X., Johnston, P., Lowe, J. K., Tian, Y., Horvath, S., Mill, J., Cantor, R. M., Blencowe, B. J. and Geschwind, D. H. 2011. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature,* 474**,** 380-4.

Warre-Cornish, K., Perfect, L., Nagy, R., Duarte, R. R. R., Reid, M. J., Raval, P., Mueller, A., Evans, A. L., Couch, A., Ghevaert, C., et al. 2020. Interferon-gamma signaling in human iPSC-derived neurons recapitulates neurodevelopmental disorder phenotypes. *Sci Adv,* 6**,** eaay9506.

Webb, B. and Sali, A. 2016. Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci,* 86**,** 2 9 1-2 9 37.

Wegiel, J., Dowjat, K., Kaczmarski, W., Kuchna, I., Nowicki, K., Frackowiak, J., Mazur Kolecka, B., Wegiel, J., Silverman, W. P., Reisberg, B., et al. 2008. The role of overexpressed DYRK1A protein in the early onset of neurofibrillary degeneration in Down syndrome. *Acta Neuropathol,* 116**,** 391-407.

Wessel, S. R., Mohni, K. N., Luzwick, J. W., Dungrawala, H. and Cortez, D. 2019. Functional Analysis of the Replication Fork Proteome Identifies BET Proteins as PCNA Regulators. *Cell Rep,* 28**,** 3497-3509 e4.

West, K. L., Kelliher, J. L., Xu, Z., An, L., Reed, M. R., Eoff, R. L., Wang, J., Huen, M. S. Y. and Leung, J. W. C. 2019. LC8/DYNLL1 is a 53BP1 effector and regulates checkpoint activation. *Nucleic Acids Res,* 47**,** 6236-6249.

White, C. L., Kessler, P. M., Dickerman, B. K., Ozato, K. and Sen, G. C. 2016. Interferon Regulatory Factor 8 (IRF8) Impairs Induction of Interferon Induced with Tetratricopeptide Repeat Motif (IFIT) Gene Family Members. *J Biol Chem,* 291**,** 13535-45.

Wisniewski, J. R., Zougman, A., Nagaraj, N. and Mann, M. 2009. Universal sample preparation method for proteome analysis. *Nat Methods,* 6**,** 359-62.

Wonders, C. P. and Anderson, S. A. 2006. The origin and specification of cortical interneurons. *Nat Rev Neurosci,* 7**,** 687-96.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C. and Micklem, G. 1995. Identification of the breast cancer susceptibility gene BRCA2. *Nature,* 378**,** 789-92.

Wu, L., Cao, J., Cai, W. L., Lang, S. M., Horton, J. R., Jansen, D. J., Liu, Z. Z., Chen, J. F., Zhang, M., Mott, B. T., et al. 2018. KDM5 histone demethylases repress immune response via suppression of STING. *PLoS Biol,* 16**,** e2006134.

Wu, W., Hill, S. E., Nathan, W. J., Paiano, J., Callen, E., Wang, D., Shinoda, K., van Wietmarschen, N., Colon-Mercado, J. M., Zong, D., et al. 2021. Neuronal enhancers are hotspots for DNA single-strand break repair. *Nature.*

Wu, Y., Li, X., Yu, J., Bjorkholm, M. and Xu, D. 2019. ASF1a inhibition induces p53-dependent growth arrest and senescence of cancer cells. *Cell Death Dis,* 10**,** 76.

Xie, J., Huang, L., Li, X., Li, H., Zhou, Y., Zhu, H., Pan, T., Kendrick, K. M. and Xu, W. 2017. Immunological cytokine profiling identifies TNF-alpha as a key molecule dysregulated in autistic children. *Oncotarget,* 8**,** 82390-82398.

Yang, S., Liu, L., Cao, C., Song, N., Wang, Y., Ma, S., Zhang, Q., Yu, N., Ding, X., Yang, F., et al. 2018. USP52 acts as a deubiquitinase and promotes histone chaperone ASF1A stabilization. *Nat Commun,* 9**,** 1285.

Yarlagadda, A., Acharya, G., Kasaraneni, J., Hampe, C. S. and Clayton, A. H. 2015. Placental Barrier and Autism Spectrum Disorders: The Role of Prolactin and Dopamine on the Developing Fetal Brain. *Innov Clin Neurosci,* 12**,** 14-7.

Yata, K. and Esashi, F. 2009. Dual role of CDKs in DNA repair: to be, or not to be. *DNA Repair (Amst),* 8**,** 6-18.

Ye, L., Zhang, T., Kang, Z., Guo, G., Sun, Y., Lin, K., Huang, Q., Shi, X., Ni, Z., Ding, N., et al. 2019. Tumor-Infiltrating Immune Cells Act as a Marker for Prognosis in Colorectal Cancer. *Front Immunol,* 10**,** 2368.

Yigit, G., Brown, K. E., Kayserili, H., Pohl, E., Caliebe, A., Zahnleiter, D., Rosser, E., Bogershausen, N., Uyguner, Z. O., Altunoglu, U., et al. 2015. Mutations in CDK5RAP2 cause Seckel syndrome. *Mol Genet Genomic Med,* 3**,** 467-80.

Young, A. M., Campbell, E., Lynch, S., Suckling, J. and Powis, S. J. 2011. Aberrant NF-kappaB expression in autism spectrum condition: a mechanism for neuroinflammation. *Front Psychiatry,* 2**,** 27.

Yu, C. E., Oshima, J., Fu, Y. H., Wijsman, E. M., Hisama, F., Alisch, R., Matthews, S., Nakura, J., Miki, T., Ouais, S., et al. 1996. Positional cloning of the Werner's syndrome gene. *Science,* 272**,** 258-62.

Zavrakidis, I., Jozwiak, K. and Hauptmann, M. 2020. Statistical analysis of longitudinal data on tumour growth in mice experiments. *Sci Rep,* 10**,** 9143.

Zeman, M. K. and Cimprich, K. A. 2014. Causes and consequences of replication stress. *Nat Cell Biol,* 16**,** 2-9.

Zemek, R. M., Chin, W. L., Nowak, A. K., Millward, M. J., Lake, R. A. and Lesterhuis, W. J. 2020. Sensitizing the Tumor Microenvironment to Immune Checkpoint Therapy. *Front Immunol,* 11**,** 223.

Zerbo, O., Qian, Y., Yoshida, C., Grether, J. K., Van de Water, J. and Croen, L. A. 2015. Maternal Infection During Pregnancy and Autism Spectrum Disorders. *J Autism Dev Disord,* 45**,** 4015-25.

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., Phatnani, H. P., Guarnieri, P., Caneda, C., Ruderisch, N., et al. 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci,* 34**,** 11929-47.

Zimmerman, A. W., Jyonouchi, H., Comi, A. M., Connors, S. L., Milstien, S., Varsou, A. and Heyes, M. P. 2005. Cerebrospinal fluid and serum markers of inflammation in autism. *Pediatr Neurol,* 33**,** 195-201.

Zou, L. and Elledge, S. J. 2003. Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science,* 300**,** 1542-8.