Identifying novel disease genes in genetically undiagnosed individuals with Rett syndrome and related neurodevelopmental disorders

Simranpreet KAUR

(ORCID: https://orcid.org/0000-0002-0689-2965)

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

March 2020

Department of Paediatrics

Faculty of Medicine, Dentistry and Health Sciences

The University of Melbourne

Brain and Mitochondrial Research Group

Murdoch Children's Research Institute

Abstract

Rett Syndrome (RTT) is a severe neurodevelopmental disorder (NDD) resulting in severe cognitive and physical impairments. Despite being predominantly caused by pathogenic variants in the methyl-CpG-binding (*MECP2*) gene, between 3 – 15% of classic and atypical RTT individuals do not have a genetic diagnosis. Classic RTT individuals exhibit an apparently normal development until 6 to 18 months of age after which developmental regression occurs. Atypical RTT individuals have many features of classic RTT but do not meet all the specific diagnostic criteria. Recently, the classification of RTT has been expanded to include individuals with clinical features overlapping RTT and other NDDs, often referred to as RTT-like individuals. The clinical and genetic diagnosis for RTT-like individuals is further complicated due to the complexity of NDDs and there is an unmet need to provide a precise genetic diagnosis for these individuals.

Next generation sequencing (NGS) studies are continuously identifying hidden genetic links between relevant molecular pathways and RTT. Through whole genome sequencing (WGS), our lab had identified a *de novo* heterozygous missense variant [c.744C>A; (p.Asp248Glu)] in the motor domain of kinesin-3 family member 1A (*KIF1A*) in one classic RTT female. Single nucleotide heterozygous variants in KIF1A have been implicated in a number of severe neurological disorders, collectively known as KIF1A-associated neurological disorders (KANDs). KIF1A encodes a neuron-specific kinesin molecular motor protein essential for ATP-dependent anterograde axonal transport of synaptic cargos along microtubules. In order to determine additional RTT/RTT-like individuals with KIF1A variations, we collected further clinical and genetic information from our collaborators of 30 individuals with 18 different missense variants affecting the critical motor domain of KIF1A. After careful clinical assessment, we identified three additional individuals with different novel missense variants exhibiting overlapping RTT-like and KAND clinical features. *In silico* tools predicted all variants to affect proper protein folding and were predicted to be likely disease causing. In addition, *in vitro* functional analysis using the highly specific neurite tip accumulation and microtubule gliding assays, demonstrated all variants to have reduced microtubule based

movement, suggesting these variants are indeed significantly pathogenic. Comparison of the clinical features of the remaining 27 KAND individuals with 16 variants in the KIF1A motor domain suggested that specific clinical features and phenotypic severity was largely dependent upon the location of the variant.

Using an NGS approach, we identified pathogenic *MECP2* variants, previously missed by mainstream genetic testing, in seven RTT individuals including a case of a mosaic male. In addition, we found variants in two genes that are known to be associated with NDDs and RTT-like syndromes; Structural Maintenance of Chromosomes 1A [*SMC1A*; c.3576delA; p.(Val1193Serfs*2)] and SH3 and multiple ankyrin repeat domains 3 (*SHANK3*; c.2265+1G>A). This work continues to expand the genetic basis of RTT.

Through whole exome sequencing (WES), we have also identified an atypical RTT female with a heterozygous nonsense variant [c.3385C>T; p.(Arg1129*)] in Lysine Acetyltransferase 6A (*KAT6A*) that encodes a chromatin remodelling protein. Heterozygous protein truncating variants in this gene have been associated with *KAT6A*-related intellectual disability. Through various collaborations, we identified a further four individuals with *KAT6A* variants [c.3820G>T; p.(Glu1274*), c.3399_3400dup; p.(Lys1134Argfs*14), c.3377delC; p.(Ser1126Phefs*8) and c.3631_3632delGT; p.(Val1211*)] who had clinical symptoms overlapping with RTT/RTT-like individuals. Through systematic re-analysis of a reported cohort of 76 individuals with *KAT6A*-related intellectual disability we recognized two additional individuals exhibiting RTT-like clinical features with *KAT6A* variants [c.4254_4257del; p.(Glu1419Trpfs*12) and c.3661G>T; p.(Glu1221*)] . All the identified variants in the seven RTT-like individuals were clustered in the last exon of *KAT6A* and *in silico* analysis predicted the variants to cause a dominant-negative effect. These seven individuals exhibited clinical features overlapping between RTT and *KAT6A*-related intellectual disability that was previously unrecognized.

Using singleton WGS, a novel heterozygous nonsense variant in another chromatin regulator gene, Chromodomain helicase DNA-binding protein 8 [*CHD8*; c.5017C>T; p.(Arg1673*)] was identified in an atypical RTT individual. Heterozygous protein truncating variants in *CHD8* have been implicated in NDDs including Autism Spectrum Disorders (ASDs). Functional analysis confirmed reduction in CHD8 protein levels in her fibroblasts, confirming the pathogenicity of the identified variant.

In another RTT-like female, a *de novo* heterozygous missense variant [c.271G>A; p.(Asp91Asn)] in the Eukaryotic Translation Elongation Factor 1 Alpha 2 (*EEF1A2*), involved in protein translation, was identified through singleton WES. This variant has been previously reported in a female with NDD. Interestingly, a single case with the same *EEF1A2* variant [c.274G>A, p.(Ala92Thr)] has also been reported in a RTT-like female. Thus, our findings further established the casual association between *EEF1A2* and a RTT-like phenotype.

In a RTT-like individual, a *de novo* large deletion at chromosome 9q34.11 (hg19:131,455,942-131,743,585) resulted in the loss of 13 genes, including the 3' end of the coding sequence of SET Nuclear Proto-Oncogene (*SET*) and the 5' end of the coding sequence of Nucleoporin-188 kDa (*NUP188*). The truncation of *SET* resulted in the loss of a highly conserved critical nucleosome assembly protein (NAP) domain crucial for assembly of nucleosomes and chromatin fluidity. Subsequent WES in the same individual identified a missense variant in *NUP188* [c.3922C>T; p.(Arg1308Cys)] on the other allele. Preliminary functional studies in individual's fibroblasts showed reduced NUP188 protein levels and enlarged nuclei, suggestive of perturbed NUP188 function. In this individual two genes, *SET* and *NUP188*, may be contributing to the affected individual's phenotype.

Interestingly, a homozygous variant in a novel candidate disease gene, Potassium Channel Tetramerization Domain Containing 16 [*KCTD16*; c.937T>A; p.(Ser313Thr)], was also revealed in a classic RTT female. KCTD16 encodes an auxiliary subunit that associates with GABA-B receptor and regulates receptor response in an agonist concentration dependent manner. Although variants in the KCTD family of proteins have previously been reported in

individuals with NDDs, defects specifically in *KCTD16* are yet to be linked with any human disease. Our preliminary electrophysiological studies in *Xenopus* oocytes investigating perturbed GABA-B receptor kinetics in response to the *KCTD16* variant [p.(Ser313Thr)] did not reveal any significant differences when compared to wild type, as well as a variant commonly found in the healthy population [p.(Asp160Asn)]. Despite this, we plan to continue these investigations in a mammalian Chinese Hamster Ovary (CHO) cell-based model to further evaluate the variant's pathogenicity.

Overall, this study has provided functional evidence of variations affecting the motor domain of KIF1A in the pathophysiology of RTT/RTT-like disorders. In addition to identifying pathogenic variations in four known RTT-related genes (*MECP2, SHANK3, SMC1A* and *EEF1A2*), in this project we have further expanded the genetic landscape of RTT/RTT-like disorders to include variations in five additional genes (*KAT6A*, *CHD8*, *SET*, *NUP188* and *KCTD16*). We recommend that the testing of these genes should be considered routinely while analysing NGS data in mutation negative RTT/ RTT-like individuals. The identification of additional members of key molecular pathways perturbed in RTT has further expanded our understanding of the underlying biology behind RTT, and this may pave the way for future targeted therapeutic options for RTT.

Declaration

This is to certify that:

(i) The thesis comprises my original work towards the PhD, except where indicated in the Preface,

(ii) Due acknowledgment has been made in the text to all materials used,

(iii) The thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

SIMRANPREET KAUR

March, 2020

Preface

Publications supporting thesis

Chapter 3

Simranpreet Kaur*, Nicole J. Van Bergen*, Kristen J. Verhey, Cameron J. Nowell, Breane Budaitis, Yang Yue, Carolyn Ellaway, Nicola Brunetti-Pierri, Gerarda Cappuccio, Irene Bruno, Lia Boyle, Vincenzo Nigro, Annalaura Torella, Tony Roscioli, Mark J. Cowley, Sean Massey, Rhea Sonawane, Matthew D. Burton, Bitten Schonewolf-Greulich, Zeynep Tümer, Wendy Chung, Wendy A. Gold^, John Christodoulou^. *Phenotypic spectrum of de novo missense variants in kinesin family member 1A (KIF1A): expansion to include Rett syndrome and Rett-like patients* (2020). Submitted to Human Mutation. Manuscript is currently under review.

* indicates co-first author; ^ indicates co-last author

Contribution to manuscript: First author publication

I was involved in all aspects of preparing this manuscript. I had conceptualised and designed the study with JC, NVB and WAG, performed 100% of the experimental work, conducted the vast majority of the data analysis and independently wrote the initial draft of this manuscript. I was primarily responsible for manuscript submission, and I co-ordinated the response to the reviewers' comments and the final manuscript preparation.

Progress of manuscript:

This manuscript was submitted to the peer reviewed journal, Human Mutation (https://onlinelibrary.wiley.com/journal/10981004) on 19th Nov 2019. The reviews from two independent reviewers, editorial board and managing editor's comments were received on 30th December 2019 (Appendix D). The revised manuscript, included as Chapter 3 in this thesis,

has been submitted to the journal on 13th Feb 2020 which is currently under review as of 9th March 2020.

Chapter 5

Simranpreet Kaur, Nicole J. Van Bergen, Bruria Ben-Zeev, Emanuela Leonardi, Tiong Y. Tan, David Coman, Benjamin Kamien, Susan M. White, Miya St John, Dean Phelan, Kristin Rigbye, Chern Lim, Michelle C. Torres, Melanie Marty, Elena Savva, Teresa Zhao, Sean Massey, Alessandra Murgia, Wendy A. Gold, John Christodoulou. *Expanding the genetic landscape of Rett syndrome to include Lysine (K) acetyltransferase 6A (KAT6A)*. (2020). Submitted to Journal of Genetics and Genomics. Manuscript is currently under review.

Contribution to manuscript: First author publication

I was involved in all aspects of this manuscript. I had conceptualised and designed the study with JC, NVB and WAG, collected clinical information from referring clinicians, conducted the majority of the data analysis and independently wrote the initial draft of this manuscript.

Progress of manuscript:

This manuscript has been submitted to the Journal of Genetics and Genomics, which is an international journal that publish peer-reviewed articles highlighting novel and significant genomic findings. The manuscript was submitted on 12th Feb 2020 and is currently under review (Appendix F).

Chapter 6

Publication 1: *Author number 9 out of total 17*

Schonewolf-Greulich, B., Bisgaard, A. M., Duno, M., Jespersgaard, C., Rokkjaer, M., Hansen, L. K., . **Kaur, S**. . . Tumer, Z. (2019). Mosaic *MECP2* variants in males with classical Rett

syndrome features, including stereotypical hand movements. Clin Genet, 95(3), 403-408. doi:10.1111/cge.13473 (Appendix G, Publication 1).

Publication 2: *Author number 6 out of total 12*

Schonewolf-Greulich, B., Bisgaard, A. M., Moller, R. S., Duno, M., Brondum-Nielsen, K., **Kaur, S**., . . . Tumer, Z. (2019). Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. Clin Genet, 95(2), 221-230. doi:10.1111/cge.13153 (Appendix G, Publication 2).

Contribution to manuscripts: Co-author publication

I collected the cohort of DNA samples for mutation-negative RTT/RTT-like individuals with variants in *MECP2* (mosaic), *SMC1A* and *SHANK3* from our collaborators. The NGS was performed using a custom designed gene panel consisting of 136 genes implicated in various NDDs. I designed the custom gene panel after carefully selecting potential genes based on a thorough review of the existing literature. Moreover, I was also involved with the raw data analysis with the first author, Dr. Bitten Schonewolf-Greulich. The full manuscripts have been included in Appendix G (Publication 1 and Publication 2).

Publication 3: *First author publication*

Kaur, S., N. J. Van Bergen, W. A. Gold, S. Eggers, S. Lunke, S. M. White, C. Ellaway and J. Christodoulou (2019). "Whole exome sequencing reveals a *de novo* missense variant in *EEF1A2* in a Rett syndrome-like patient." Clin Case Rep 7(12): 2476-2482.

Contribution to manuscript:

I conceptualised and designed this study with JC, NVB and WAG, collected clinical information and DNA samples from referring clinicians, conducted the majority of the data analysis and wrote the initial draft of this manuscript. I was primarily responsible for

manuscript submission and addressing reviewer's comments. Thus, I was involved in all aspects of this manuscript.

Appendix G

Kaur, S., Christodoulou, J. *MECP2* Disorders. 2001 Oct 3 [Updated 2019 Sep 19]. In: Adamm M.P., Ardinger, H.H., Pagon, R.A., et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1497/

Contribution to book chapter: First author

GeneReviews is an online database published as a chapter on the National Center for Biotechnology Information Bookshelf site. It contains standardized peer-reviewed articles describing information regarding Mendelian genetic disorders, critical for diagnosis and management of affected individual and genetic counselling of the involved families. I was mainly involved in the revision of the prepared draft of this article. I have also added ~20% additional information in this article through a primary literature search. The full manuscripts have been included in Appendix G (Book Chapter 1).

Acknowledgements

To Prof John Christodoulou (*my primary supervisor*) – It has been a privilege to do my PhD with you John, as you are an excellent supervisor! Thank you so much for giving me this fantastic opportunity to work with you again after my M.Phil. (Medicine). I really appreciate each and every word of your appreciation, support and encouragement throughout my PhD journey. You gave me wings to fly and explore the path of PhD under your constant guidance. I will always adore you for guiding me to handle Ekam's (my 3 year old son) health issues. This PhD would not have been possible without the flexibility that you had provided me. Last but not the least, thank you so much for giving top up scholarship and letter of support for various other scholarships that helped me to survive during my struggling times in life. Through your continuous support, I was able to manage PhD and parenting at the same time without any relative's backing.

To Dr. Nicole Van Bergen (*my associate supervisor*) – My sincerest thanks to you Nic for teaching me different skills to be a better scientist. Your attention to details, great writing skills, and ability to handle different projects always amaze me. Thank you so much for giving me an opportunity to deal with my PhD sub- projects independently! You have been a great mentor to me. I really appreciate all your time to guide me through this journey.

To Dr. Wendy Gold (*my associate supervisor*) – Thank you so much for taking me as your student once again! It's been a dream come true to do my PhD studies with you and John years after passing my M.Phil. studies. You have been a great inspiration to me. I really admire the way you had taught me lab skills years ago that I still apply and share with other researchers. You have always been my mentor and one of my strongest support pillars. Your confidence in my skills made me a great scientist. Thank you so much Wendy for listening to my personal issues and guiding me through. I consider myself lucky to have known you.

To Dr. Shireen Lamande and A/Prof Sue White (*my PhD advisory committee members*) – I would like to express my gratitude to you for your invaluable advice and words of encouragement during my PhD studies. Thank you for your support and keeping me on track for completing my studies in time.

To Prof Kristen Verhey (*my KIF1A collaborator*) – I am really thankful to you for sharing your immensely valuable experience in the KIF1A research. I was really touched with your warm welcome when I had visited your lab in Michigan, USA. I sincerely thank all of your team members, especially Ms Breane Budaitis and Yue Yang, to teach me critical skills for KIF1A research. Under your kind guidance, I was able to set up different KIF1A based research assays in our lab.

To Prof Wendy Chung and Ms Lia Boyle (*my KIF1A collaborator*) – I really admire your tireless efforts to support kids with *KIF1A* defects all throughout the world. Thank you so much for sharing clinical information of *KIF1A* individuals in your cohort with us for my PhD studies. It was a great opportunity to have met you during ASHG2019 conference in Houston. I hope that our path cross again in future.

To all affected individuals, families and other collaborators – My sincerest thanks to all affected individuals, their families and consulting physicians who participated in my PhD studies. I am truly blessed to have contributed towards the diagnosis of a number of genetically undiagnosed individuals with Rett and Rett-like syndrome.

To The University of Melbourne, Graduate Woman Victoria, Dr. Jenneth Sasse and family and Franklin Women – Thank you for supporting and funding my research. I had survived financial hiccups only through your kind support.

To Murdoch Children's Research Institute, Australian Functional Genomics, Human Genetics Society of Australasia Australasian Society for Stem Cell Research, and The National Stem Cell Foundation of Australia – I sincerely thank you for supporting my conference travels and lab visits. Without your support, I would not be able to get the incredible experience of sharing and learning new skills.

To my past and present lab members of Brain and Mitochondrial Group -

Prof David Thorburn- Thanks David for all your support during my PhD studies! I have learned a lot from you. You active interest in my project definitely made it better.

Nicole Andrew – Thanks Nicole for helping me around! I will always remember your words "Simran, be kind to yourself". You always took time to listen to me and guide me through tough times.

Sean – Thank you so much for all your help with the lab work. I will miss our quick chats in the lab. You are one of the most hard working lab scientist that I have ever met. Rise and Shine in your career!

Simone, Alison and Ann –Thank you so much for encouraging me when I felt down. I truly admire your skills in balancing personal and professional life. All your hugs mattered a lot to me. Also, I learned a lot of tips from you regarding Ekam and thank you so much for sharing little gifts along the way!

Rocio, Nat, Yau, Tegan, AnneMarie and Cameron – Thanks so much to each one of you to help me whenever I got stuck! You have been really kind in helping me around. I am sorry that I couldn't go out for a lot of coffee breaks (just a limitation of a mum handling different things at one time). I am sure you understand!

Shalini, Marilou and Marzena – Thanks you so much girls for listening to me when I needed to vent! Our talks made me cheerful! I have learned a lot from each one of you.

To my family

To Sant Wareyam Singh Ji, Sant Lakhbir Singh Ji and all my sangat family – Thank you so much for all your love, support and blessings. I am indebted to all your teachings to keep me on track during my PhD studies. This journey would be impossible without your constant guidance.

To Late Harbans Kaur (*my mother*) - Mumma - This PhD is for you! I have finally fulfilled your one more dream. Life has not been easy without you and I miss you with each and every breath of mine. I consider myself extremely lucky to have you as my mother. You sacrificed a lot in your life to make me independent. I will always be indebted for all your teachings. You vision to support girl's education has made a significant change in my life. I do not have words to thank you for everything that you had done for me!

To Mukhtiar Singh (*my father*) – Thanks Dad for all your unconditional love and financial support! Its takes real courage to send a daughter alone in a foreign country and spend thousands of dollars on her education, especially in community from where we come from. I am so proud of you that you broke that stereotype and supported me to pursue my dreams. I still remember that you always told me "Beta, you are my lucky charm and I am there for you always!". Love you Dad and I hope that I have made you proud!

To Ekam – I am truly blessed to have you in my life. Although you are unaware, but I want to tell you that you gave me a lot of strength in my challenging times. Your one smile would fade away all my fears. I am sorry for not being there with you at times when you needed me. Thank you so much for coming along with me to different lab meetings, to my office area while I had to quickly do something in lab and for listening to all my research talks. I want you to remember that you have contributed a lot in your Mum's PhD.

To Sukhpreet Singh (*my husband*) – Thank you for being there for Ekam whenever I couldn't be around for him. You were always so flexible in changing your job timings so that I could come to work early. I think we have done a good job in managing Ekam's health with our respective careers. Thank you so much for all your help and support!

To Kaminee, Kanchan, Jaspreet and Nitin – Although you are miles away, but you have always supported me emotionally endlessly during this journey. I always got a new perspective from you all towards ongoing situation. Thank you so much for your presence in my life! Blessed to have you all... Kaminee - I cannt repay you ever for what you have done for my mum when I was unable to be there for her!!

To Kinder Di, Babby Veerji and Jazlyn – Thank you so much for being there with us when we needed you the most. You all took care of Ekam and Sukhpreet so well when I was away from home attending different conferences. Kinder Di - You were only a phone call away when I needed you and I will always cherish our conversations. I will always be indebted for all your time and efforts in keeping us together.

To Mandeep Bhabhi ji, Vikram, Sukham and Balraj – Thank you all for your support!

Table of Contents

List of Figures

List of Tables

Abbreviations

%	Percentage
º	Degrees
ACMG	The American College of Medical Genetics and Genomics
ADHD	Attention deficit hypersensitivity disorders
ADP	Adenine diphosphate
ASD	Autism spectrum disorder
ATP	Adenine triphosphate
BDNF	Brain Derived Neurotrophic Factor
bp	Base pairs
BSA	Bovine serum albumin
C	Celsius
Cat. no.	Catalogue number
CDKL5	Cyclin-dependent kinase-like 5
CGH	Comparative genomic hybridization
CHD8	Chromodomain helicase DNA-binding protein 8
CHO	Chinese Hamster Ovary
CNS	Central Nervous System
CO_2	Carbon-dioxide
COS	CV-1 in Origin with SV40 genes
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
dNTPs	Deoxy ribonucleotide triphosphates
E coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
EEF1A2	Eukaryotic Translation Elongation Factor 1 Alpha 2
FBD	Forkhead-binding domain
FBS	Fetal Bovine Serum
FHA	Fork-head domain
FOXG1	Forkhead box protein G1
g	Gram
GABA	gamma-Aminobutyric acid
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GIRK	G protein-coupled inwardly-rectifying potassium channel
gnomAD	Genome Aggregation Database
GTP	Guanosine triphosphate
HDAC	Histone deacetylases
HSNIIC	Hereditary sensory neuropathy type IIC
IMPDH2	Inosine Monophosphate Dehydrogenase 2
KAND	KIF1A-associated neurological disorder

KAT6A	Lysine acetyltransferase 6A
kb	kilo-base pair
KCTD	Potassium Channel Tetramerization Domain
KCTD16	Potassium Channel Tetramerization Domain Containing 16
KIF	kinesin superfamily
KIF1A	kinesin family member 1A
KIF23	Kinesin Family Member 23
L	Litres
LB	Lysogeny broth
LOVD	Leiden Open (source) Variation Database
LZ	Leucine zipper
M	Molar
MBD	methyl-CpG-binding domain
mCit	mCitrine
MCRI	Murdoch Children's Research Institute
MECP2	Methyl-CpG-binding protein 2
MFI	Mean florescence intensity
$MgCl_2$	Magnesium chloride
Min	Minute
mNG	mNeonGreen
MRD9	Non-syndromic intellectual disability 9
MRI	Magnetic resonance imaging
msec	milli-second
NC	Neck coil
NCBI	National Center for Biotechnology Information
NDD	Neurodevelopmental disorders
NGS	Next generation sequencing
NGS	Next Generation Sequencing
NLS	Nuclear localization signal
NUP188	Nucleoporin-188 kDa
OMIM	Online Mendelian Inheritance in Man
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PEHO	Progressive encephalopathy with Edema, Hypsarrhythmia and Optic atrophy
PFA	Paraformaldehyde
PMSF	Phenylmethylsulfonyl fluoride
PTM	Post translational modifications
PVDF	Polyvinylidene fluoride
RA	Retinoic acid
rpm	Rotations per minute
RPMI	Roswell Park Memorial Institute

RTT	Rett syndrome
sec	Seconds
SD	Standard Deviation
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis
SEM	Standard Error of the Mean
SPG30	Hereditary spastic paraplegia 30
TBE	Tris-borate-EDTA
TIRF	Total internal reflection microscopy
T_m	Melting temperature
TRD	Transcription repression domain
U	Unit
UTR	Untranslated region
V	Volts
VCGS	Victorian Clinical Genetic Services
VCGS	Victorian Clinical Genetics Service
Vol	Volume
WES	Whole Exome Sequencing
wg	Weeks of gestation
WGS	Whole Genome Sequencing
wt/vol	Weight/ volume
µ	Micro
µg	Micrograms
µl	Microlitres
µM	Micromolar

Chapter 1

Chapter 1 Introduction and Literature review

1.1 Neurodevelopmental disorders

Brain development is a well-timed regulated process such that any deviations from the normal developmental trajectory results in neurodevelopmental disorders (NDDs). NDDs constitute a large group of common and complex disorders resulting in heterogeneous phenotypes but mainly affecting social behaviour, motor, language and cognitive abilities of affected individuals at varying severity levels during development (Cardoso et al., 2019). Various genetic, epigenetic and environmental factors contribute to the clinical and behavioural phenotypes observed in affected individuals, thus further complicating the understanding the complex mechanisms behind the pathogenesis of NDDs (Mullin et al., 2013). The multifaceted wide spectrum of NDDs includes, but is not limited to, intellectual disability, autism spectrum disorder (ASD), attention deficit hypersensitivity disorders (ADHD), speech disorders, Tourette syndrome, epileptic disorders, learning disability, Down syndrome, Angelman syndrome, Prader-Willi syndrome, Fragile X syndrome and Rett syndrome (RTT). Often affected individuals exhibit phenotype overlapping with a number of NDD disorders, thus the genotype-phenotype correlations of these NDDs is constantly expanding (Cardoso et al., 2019). Dysregulation of common molecular pathways further increases the complexity behind the crosstalk between different multifactorial NDDs (Kumar et al., 2019). In addition, individuals affected by some NDDs often display many, but not all of required clinical diagnostic criteria for a particular disorder, and thus there is constant expansion of the phenotype of NDD-like individuals.

Owing to their complex nature, the majority of NDDs are polygenic, with only handful of NDDs having monogenic cause, including RTT, which is one of the leading causes of severe intellectual disability and developmental regression in females.

1.2 Rett syndrome (RTT)

RTT (OMIM 312750), first reported by Andreas Rett in 1966, is a severe, X-linked dominant postnatal progressive neurodevelopmental disorder that primarily affects females (Hagberg, Aicardi, Dias, & Ramos, 1983; Rett, 1966), with the worldwide incidence ranging between 1 in 10,000 to 23,000 live female births (A. H. Armstrong, Hangauer, Agazzi, Nunez, & Gieron-Korthals, 2010). An Australian-epidemiological study suggested that 1 in 8500 females in Australia are diagnosed with RTT by the age of 15 years, making it the most common cause of severe intellectual disability in females after Down Syndrome (Fehr et al., 2011). Before establishing the genetic cause, affected individuals are often first clinically classified as per diagnostic criteria established for RTT (Neul et al., 2010).

1.3 Clinical Classification of RTT

In 1988, International Rett syndrome Association and the Centres for Disease Control outlined key criteria's for the clinical diagnosis of RTT; however these diagnostic criteria were further simplified by the RettSearch Consortium, which is now known as Neul's revised diagnostic criteria for RTT (Neul et al., 2010). It includes four main, two exclusion and eleven supportive criteria as outlined in Table 1.1. Based on the criteria, an affected individual can be classified broadly into two main categories; classic (or typical) and atypical (or variant) RTT (Neul et al., 2010).

1.3.1 Classic (or typical) RTT

The key requirement for the clinical diagnosis of classic RTT is the evidence for a period of regression followed by recovery or stabilization. In addition, an individual must fulfil all four main and both the exclusion criteria outlined in Neul's revised diagnostic criteria (Table 1.1) (Neul et al., 2010). Although supportive criteria fulfilment is not necessary, most individuals exhibit clinical symptoms as outlined in the supportive criteria.

Table 1.1: The revised diagnostic criteria for RTT from (Neul et al., 2010).

Criteria	Classic/ typical	Atypical/ variant RTT
Essential criteria (1) Regression followed by recovery	Yes	Yes
Main criteria (4) 1. Partial/ complete loss of acquired purposeful hand skills 2. Partial/ complete loss of acquired spoken language 3. Gait abnormalities 4. Stereotypic hand movements	4	≥ 2
Supportive criteria (11) 1. Breathing disturbances when awake 2. Bruxism when awake 3. Impaired sleep pattern 4. Abnormal muscle tone 5. Peripheral vasomotor disturbances 6. Scoliosis/kyphosis 7. Growth retardation 8. Small cold hands and feet 9. Inappropriate laughing/screaming spells 10. Diminished response to pain 11. Intense eye communication	Not required (although often present)	≥ 5
Exclusion criteria (2) 1. Brain injury 2. Grossly abnormal psychomotor development (first 6 months)	2	0

Classic RTT individuals follow the similar course of disease progression which is divided into four main stages, however the age of onset and its duration has been found to be variable among individuals. During **stage 1**, the infant's physical and neurological development progresses apparently normally until 6-18 months of age (Figure 1.1). Individuals may exhibit deceleration of head growth (acquired microcephaly) and hypotonia. Following a brief period of developmental stagnation, there is a rapid regression phase, **stage II**, in which loss of attained milestones, particularly speech and fine or gross motor skills occurs. In addition, various autistic features including lack of eye contact, loss of social interaction and limited facial expression become more prominent as the affected individual becomes older (Nomura, 2005). Moreover, affected individual develop deceleration of head growth (usually noticed in this stage) and characteristic stereotypic hand movements (wringing and washing), which is a hallmark of classic RTT. During **stage III**, the clinical phenotype may stabilize or an individual may develop bruxism, screaming/ crying/ laughing fits, severe epileptic seizures, breathing irregularities and intense eye gaze. During **stage IV,** affected individuals lose ambulation and experience muscle rigidity, spasticity, weakness, dystonia and scoliosis. Moreover, the

stereotypic hand movements become less intense, and eye contact along with communication, remain intact (Jian et al., 2006; Williamson & Christodoulou, 2006). Approximately a quarter of deaths of RTT individuals are sudden and may be due to cardiac abnormalities (De Felice et al., 2012).

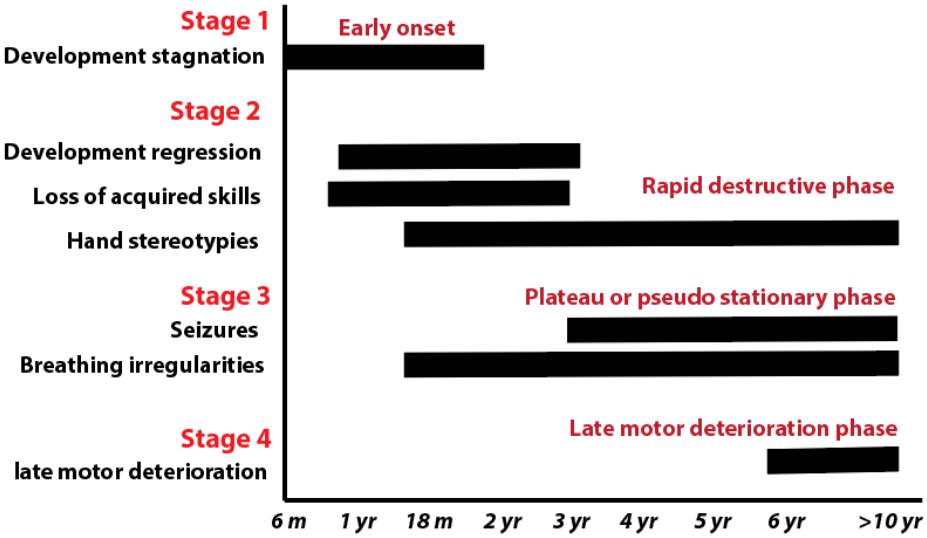


Figure 1.1 The various clinical stages of classic RTT.

The clinical progression of classic RTT is divided into four stages. After a normal development between 6-18 months of age, the development of RTT individual halts during stage 1, and loss of acquired skills along with stereotypic hand movements occur during stage 2. Breathing irregularities and seizures become evident during stage 3 whereas late motor deterioration manifests during the last stage, stage 4**.**

1.3.2 Atypical (or variant) RTT

Affected individuals who exhibit many clinical features of RTT; however but do not yet fulfil all of the main criteria to be classified as classic RTT, are referred to as atypical RTT (Oexle, Thamm-Mucke, Mayer, & Tinschert, 2005; Zappella, Meloni, Longo, Hayek, & Renieri, 2001). An atypical RTT clinical diagnosis is based on a period of regression, followed by stabilization or partial recovery, along with the attainment of two main and at least five supportive criteria's according to the revised Neul diagnostic criteria (Neul et al., 2010). Among the broader clinical spectrum of atypical RTT individuals, different forms have been described in the literature exhibiting either much milder or more severe clinical features (Table 1.2) as compared with classic RTT (Kalscheuer et al., 2003; Neul, 2012; Pini et al., 2012; Zappella et al., 2001).

Table 1.2: List of subtypes of atypical RTT with characteristic clinical features.

Subtypes of atypical RTT	Characteristic clinical features	Reference
'forme fruste' variant	Milder version of classic RTT, late regression phase, less severe stereotypic hand movements	(Hagberg & Rasmussen, 1986)
Rolando variant (Congenital variant)	Infantile hypotonia, abnormal development from birth, developmental defects during first three months of life, characteristic muscle wasting, absence of eye gaze, breathing difficulties, lack of mobility	(Ariani et al., 2008)
Zappella variant (preserved speech variant)	Regression at 1-3 years, prolonged plateau phase, better retained hand use, language recovery after regression, milder intellectual disability, autistic behaviour, decreased frequency of typical RTT features, rarely epilepsy, rare autonomic dysfunction, milder scoliosis/ kyphosis, normal head circumference; normal height and weight	(De Bona et al., 2000; Marschik, Einspieler, Oberle, Laccone, & Prechtl, 2009; Stembalska, Gil, & Pesz, 2011; Townend et al., 2015)
Hanefeld variant (Early onset variant or infantile seizure onset variant)	Onset of seizures occurs before 6 months of age followed by other classic RTT features.	(Hanefeld, 1985; Scala et al., 2005)

While RTT predominantly affects females, up to 60 male cases with RTT have been listed in RettBASE (http://mecp2.chw.edu.au/mecp2/; curation database to collate variants in genes associated with RTT) (Krishnaraj, Ho, & Christodoulou, 2017). Severe neonatal-onset encephalopathy is often observed in male RTT cases, which is characterised by severe seizures, breathing abnormalities, abnormal muscle tone, a metabolic-degenerative type of clinical pattern, involuntary movements, and death often occurs before two years of age. There have been eight published reports of males exhibiting classic RTT features and with somatic mosaicism of RTT-related gene expression (Clayton-Smith, Watson, Ramsden, & Black, 2000; Schonewolf-Greulich, Bisgaard, Duno, et al., 2019; Topcu et al., 2002).

More recently, a new clinical sub-classification, known as RTT-like, has been suggested, that represents a large proportion of individuals who exhibit a number of clinical features of RTT, but not enough in order to be classified as either classic or atypical RTT. Even though there is no formal definition of RTT-like cases, an increasing number of genes are being associated with clinical phenotypes overlapping with RTT features, along with features characteristic of other NDDs, and so the genetic landscape of RTT and RTT-like individuals is constantly evolving (Iwama et al., 2019).

1.4 Neuropathology of RTT

Several histopathological studies of RTT individual autopsy brain samples revealed consistent features in the structure and function of neurons and synapses.

Neuropathological post-mortem studies of RTT individuals revealed smaller brain size in RTT individuals when compared to age- and height- matched healthy individuals, which is consistent with the clinical presentation of acquired microcephaly (Figure 1.2) (Reiss et al., 1993). No continual decrease in brain size with age was observed, which suggests RTT is not a neurodegenerative disorder (D. D. Armstrong, 2002; D. D. Armstrong et al., 1999). The reduction in brain weight was not found to be generalized, but rather region specific. Reduced brain volumes were observed in prefrontal, posterior frontal, and anterior temporal regions, whilst posterior temporal and posterior occipital regions remained unaffected (Reiss et al., 1993; Subramaniam, Naidu, & Reiss, 1997).

The cerebral cortex, thalamus, basal ganglia, amygdala, and hippocampus of RTT individuals showed generalized reduced neuronal size with increased cell-packing density and no indication of active degeneration Figure 1.2 (Bauman, Kemper, & Arin, 1995). Reduced numbers of neurons were observed mainly in frontal and temporal cortex regions, particularly in layers II and III (Belichenko, Oldfors, Hagberg, & Dahlstrom, 1994). In addition, a significant decrease in the dendritic arborisation was also observed in the frontal, motor, and subicular cortices suggesting growth and developmental arrest further reinforcing the impaired neurodevelopmental basis of RTT (D. D. Armstrong et al., 1999).

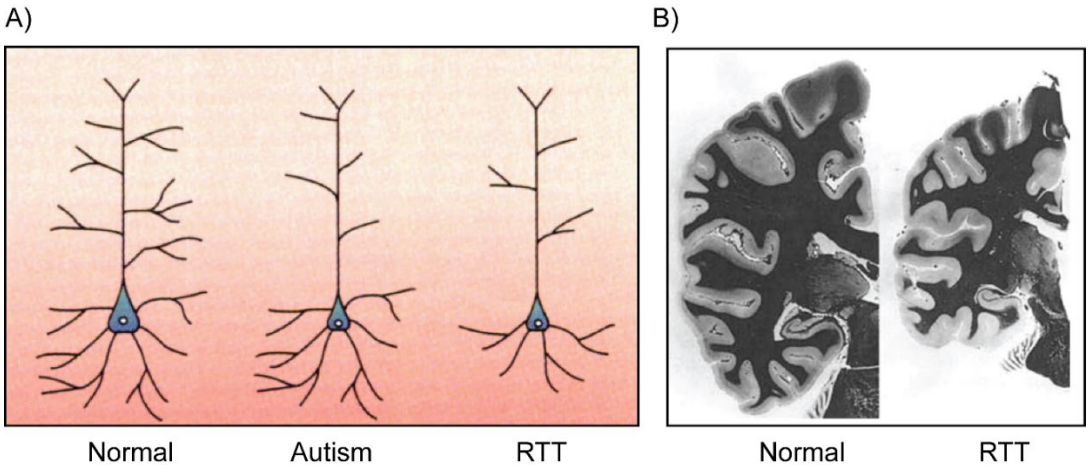


Figure 1.2: Neuronal and brain abnormalities in RTT individuals.

A) Reduced cell body and less complex dendritic branching of pyramidal neurons in individuals with autism or RTT individual compared to controls from (Zoghbi, 2003). B) Cerebral hemisphere coronal section from a 21-years old age matched healthy individual (left) and RTT individual (right) showing overall reduced brain size in the RTT individual.

1.5 Genetic causes of RTT

The genetic basis for RTT was uncovered by Amir et al in 1999, when *de novo* heterozygous pathogenic variants in the X-linked methyl-CpG-binding protein 2 (*MECP2*; OMIM: 300005) gene were identified in several classic RTT individuals (Amir et al., 1999). This discovery provided a major breakthrough towards understanding the genetics of this severe NDD. Since 1999, a total of 4,668 variants (including pathogenic, benign or variants of unknown significance have been reported in RettBASE, which are dispersed throughout the MeCP2 protein (Figure 1.3) (Krishnaraj et al., 2017). The most recurrent pathogenic missense and/nonsense variants in *MECP2* affects key functional domains including the methyl binding domain [MBD; p.(Arg106Trp), p.(Arg133Cys), p.(Thr158Met) and p.(Arg168*)] and transcription repression domain [TRD; p.(Arg255*), p.(Arg270*), p.(Arg294*) and p.(Arg306Cys)], which together account for about 39% of reported *MECP2* variants (Krishnaraj et al., 2017; Zachariah & Rastegar, 2012). These pathogenic variants are primarily cytosine to thymine transitions in CpG islands.

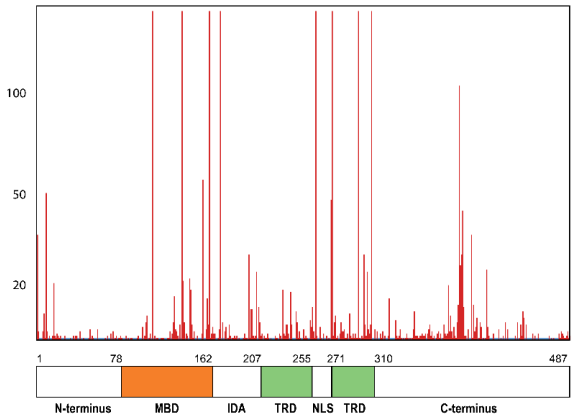


Figure 1.3: Mutation spectrum in *MECP2*.

Distribution of commonly observed variants (red lines) relative to the MeCP2B domain structure (1 to 487 amino acid). MBD = methyl-CpG binding domain, TRD = transcription repression domain, NLS = nuclear localization signal. Figure adapted from (Krishnaraj et al., 2017).

1.5.1 Methyl CpG Protein 2

The methyl-CpG-binding domain (MBD) family of proteins specifically binds to methylated sequences within the genome. They are critically involved in crosstalk between DNA methylation, histone modifications and chromatin remodelling for transcriptional regulation (Du, Luu, Stirzaker, & Clark, 2015). MeCP2, a 53kDa nuclear protein, was the first MBD-containing protein discovered with a multifaceted role as a gene expression activator, repressor, splicing regulator and long range chromatin remodelling (Ausio, Martinez de Paz, & Esteller, 2014).

In mammals, the alternate splicing of the *MECP2* gene results in two transcripts; *MECP2E1* and *MECP2E2*. The longer transcript *MECP2E1* (or *MECP2B* or *MECP2α*) comprising of exons 1, 3, and 4 and encodes for MeCP2_e1 (or MeCP2B; 498 amino acids) protein, with and the second shorter transcript *MECP2E2* (or *MeCP2A* or *MECP2β*) that consists of exons 2, 3, and 4, and encodes for the MeCP2_e2 (or MeCP2A; 486 amino acids) protein (Figure 1.4) (Ausio et al., 2014; Kriaucionis & Bird, 2004). Overall, MeCP2 consists of four main functional domains: *Methyl-CpG-binding domain* (MBD; 85 amino acid) which is encoded within exon 3 and 4 and allows MeCP2 to bind with single symmetrical methylated CpG sites with high affinity, the *transcriptional repression domain* (TRD; 104 amino acids) that is

primarily encoded by exon 4 and interacts with the transcriptional co-repressor Sin3A for recruitment of histone deacetylases (HDAC); HDAC1 and HDAC2, the *Nuclear Localization Signal* (NLS), which is critical for nuclear localisation of MeCP2, and a *C-terminal WW domain* that interacts with splicing and transcription factors (Buschdorf & Stratling, 2004).

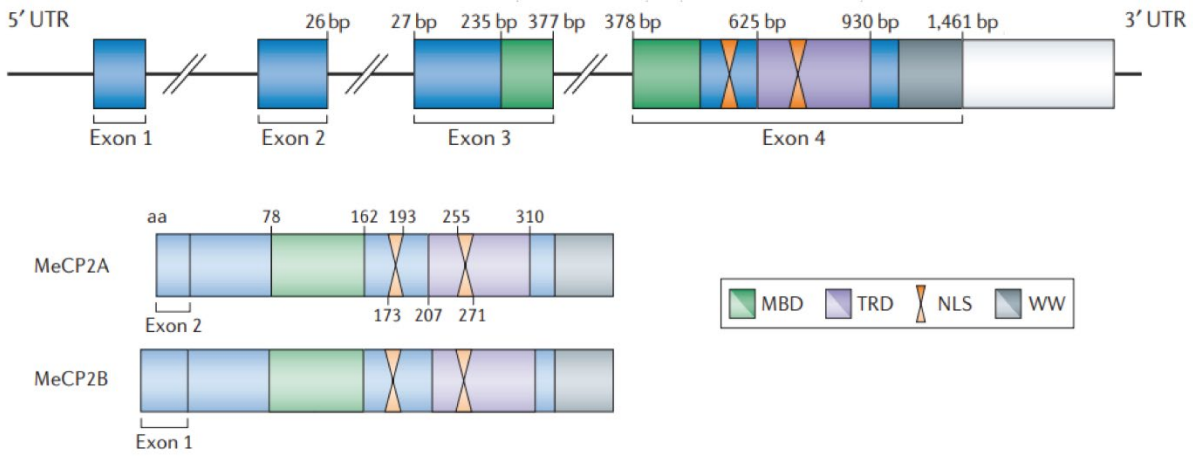


Figure 1.4: Gene and protein structure of *MECP2*.

A) Gene structure of *MECP2* labelled with exon number (below) and base pair boundaries (above). B) Schematic showing the two main isoforms of MeCP2: MeCP2A (486 amino acid) and MeCP2B (498 amino acid) that mainly differ at their N-terminal region. While in case of MeCP2A, exon 2 codes for the N-terminus, however exon 1 forms it in case of MeCP2B. MBD = methyl-CpG binding domain; NLS = nuclear localization signal, TRD = transcriptional repression domain. Figure adapted from (Bienvenu & Chelly, 2006).

MeCP2 is enriched with a variety of post translational modifications (PTMs) including acetylation, phosphorylation, ubiquitination and sumoylation, enabling MeCP2 to interact with various proteins to facilitate controlled transcriptional regulation of gene expression (Cheng et al., 2014; Cohen et al., 2011; Gonzales, Adams, Dunaway, & LaSalle, 2012; Tao et al., 2009; Zhou et al., 2006; Zocchi & Sassone-Corsi, 2012). Moreover, MeCP2 levels are spatially and developmentally regulated in all tissues but particularly high levels have been reported in mature neurons in the brain (Nagarajan, Hogart, Gwye, Martin, & LaSalle, 2006). Although the expression of both isoforms of *MECP2* is ubiquitous, the expression levels are tissue specific and are developmentally regulated. While *MECP2E2* is predominantly expressed in fibroblasts and, lymphoblast cells, *MECP2E1* expression is highest in brain, thymus, lung and developing neurons (Kriaucionis & Bird, 2004). MeCP2 expression is low during embryogenesis and

progressively increases with neuronal maturation in the postnatal period (Kishi & Macklis, 2004; Skene et al., 2010). MeCP2 expression in human CNS initiates at the brainstem nuclei and cerebral cortex Cajal-Retzius neurons at 10 week of gestation (wg) followed by increased expression in the cerebellum, mid brain, thalamus and deep cortical layers at the end of 26 wg. Around 35wg MeCP2 expression becomes noticeable in the hypothalamus, basal ganglia, hippocampus and the superficial cortical layers (Figure 1.5) (Shahbazian, Antalffy, Armstrong, & Zoghbi, 2002).

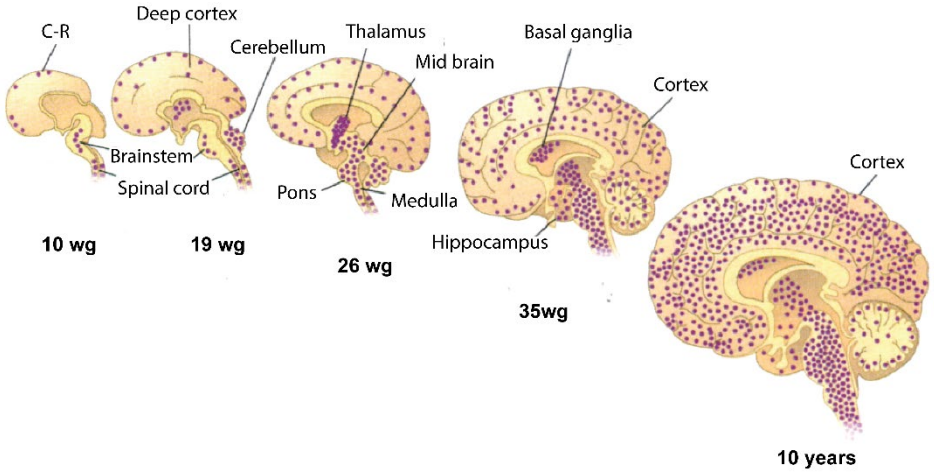


Figure 1.5: MeCP2 spatial and temporal distribution during human development.

MeCP2 is first expressed in spinal cord and last in cortex consistent to the ontogeny of the central nervous system. Cajal-Retzius (C-R) neurons express MeCP2 first followed by midbrain, thalamus, cerebellum and deep cortical neurons. wg=weeks of gestation. Image taken from (Zoghbi, 2003).

1.5.2 Other genes already implicated in RTT

Approximately 97% of classic RTT individuals, who satisfy all the main clinical criteria of RTT, carry pathogenic variants in the *MECP2* gene. On the other hand, only up to 86% of atypical RTT individuals with several necessary Neul's diagnostic criteria carry *MECP2* variants (Neul et al., 2014). A minor proportion of RTT individuals have also been found to carry pathogenic variants in other genes such as cyclin-dependent kinase-like 5 (*CDKL5*) and Forkhead box protein G1 (*FOXG1*).

CDKL5 is a 118kDa serine/threonine kinase that is critical for the development and function of the brain. CDKL5 is involved in neuronal migration, axonal growth, dendritic proliferation and synapse development. CDKL5 is highly expressed in brain particularly in the nuclei and dendrites of neurons, with expression peaking during early post-natal life (Olson et al., 2019). Variants in *CDKL5* have been reported in 498 RTT individuals in RettBase (Krishnaraj et al., 2017). CDKL5 deficiency causes severe early onset seizures, intellectual disability, progressive microcephaly and subtle facial dysmorphism (Evans et al., 2005). In addition, individuals may later exhibit stereotypic hand movements, breath holding or hyperventilation episodes, gait abnormalities, limited hand use and severe speech delay. In the past, individuals with variants in *CDKL5* were classified as "Hanefeld" or early onset epilepsy variant of RTT (Neul et al., 2010), however, more recently, these individuals are assigned into a separate clinical category known as CDKL5 deficiency disorder which is now recognised as a clinically distinct disorder from RTT (Fehr et al., 2013).

FOXG1 is a DNA binding transcription factor with a well conserved forkhead-binding domain (FBD). FOXG1 is critical for forebrain development, and is mainly involved in maintaining the balance between neuronal differentiation, progenitor proliferation and migration and integration of pyramidal neurons into the cortical plate during forebrain development (Miyoshi & Fishell, 2012). Moreover, FOXG1 directly interacts with the *MECP2E2* isoform preventing cerebellar neuronal apoptosis (Dastidar, Narayanan, Stifani, & D'Mello, 2012). Genotype-phenotype studies revealed severe phenotypes in individuals with FOXG1 variants were associated with protein truncating variants in the N-terminal region and FBD domain, whereas milder phenotypes were correlated with missense variants in FBD (Mitter et al., 2018; Vegas et al., 2018). The less frequent identification of variants in *FOXG1* in RTT individuals, classified as congenital variant of RTT, has highlighted the primary role of autosomal genes in RTT, and further challenges the notion that RTT is only an X-linked disorder (Ariani et al., 2008; Zhang et al., 2017). Recently, individuals with variants in *FOXG1* have been classified with *FOXG1* syndrome, which is now clinically independent to RTT. Individuals are characterised with profound motor deficits, abnormal breathing, stereotypic hand movements, lack of speech, poor eye contact, seizures, progressive microcephaly and an abnormal MRI with myelination defects and corpus callosum abnormalities (Kortum et al., 2011; Mitter et al., 2018).

1.5.3 Expanding genetic spectrum of RTT

In some *MECP2* mutation-negative RTT cases, the clinical symptoms do not belong distinctly to either classic or atypical RTT, and may be classified as RTT-like individuals. Moreover, there is considerable overlap of their clinical features with other NDDs that often results in difficulty with obtaining a precise diagnosis. Recently, next generation sequencing (NGS) approaches combined with array techniques (e.g. Comparative genomic hybridization; aCGH) have been used to screen for novel genes that are associated with the RTT-like symptoms in such individuals (Appendix A, List A1).

Lopes et al (2017) recently investigated a cohort of 19 Portuguese mutation-negative RTT-like individuals and determined pathogenic or very likely pathogenic variants in 13 individuals. Of these, 6 individuals had variants in genes associated with NDDs; *EEF1A2, STXBP1, ZNF238, SLC35A2, ZFX, SHROOM4* and *EIF2B2*. Moreover, they found candidate variants in five additional novel genes; *RHOBTB2, SMARCA1, GABBR2, EIF4G1* and *HTT* (Figure 1.6) (Lopes et al., 2016). Interestingly, network analysis revealed that all of these genes interact with each other at a protein level and with other known RTT associated gene products (MeCP2, CDKL5 and FOXG1) implying a potential genotype-phenotype overlap between NDD and RTT.

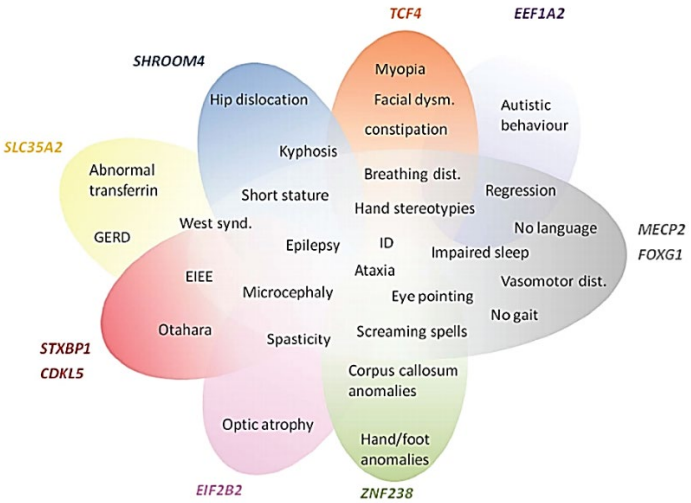


Figure 1.6 Genotype-phenotype interactions between RTT and other NDD.

Schematic representing the phenotypic overlap of genes related with RTT and other NDDs. Image taken from (Lopes et al., 2016).

In the same year, Neul et al (2017) also identified pathogenic variants in genes previously known to cause distinct NDDs in a cohort of 22 mutation-negative RTT individuals with clinical diagnosis of classic RTT (n=11) and atypical RTT (n=11). Three of these individuals were found to have pathogenic variants in *MECP2* that were previously missed during their clinical diagnosis. Variants in a total of 46 genes previously implicated in NDDs were identified in this cohort. Interestingly, 10 of the 19 undiagnosed cases had more than one likely pathogenic variant which suggests that more than one gene could contribute to the RTT-like phenotype. Moreover, one of the confirmed cases with a *MECP2* mutation had an additional five *de novo* variants, further emphasising that a high burden of cumulative mutations could be a novel feature of RTT. Furthermore, the identified variants were present in known chromatin regulators (*ACTL6B, BRD1, CHD4, HDAC1, SMARCB1* and *TRRAP*). Four genes involved in glutamate receptor signalling were also enriched in RTT-like individuals (*GABRB2, GRIN2A, GRIN2B,* and *SHANK3*), which may dysregulate neuronal excitation and contribute to seizures in individuals. The authors suggested that the aetiology of mutation-negative RTT may be genetically heterogeneous, as no variant was found to occur more than once (Sajan et al., 2017).

Whole exome sequencing (WES) carried out by Lee et al in 2016 reported the compound heterozygous variants in *ST3GAL5* [p.(Gly201Arg) and p.(Cys195Ser)], encoding GM3 synthase, in two mutation-negative Korean female siblings presented with RTT-like phenotype. GM3 synthase deficiency causing the neurological disorder was further confirmed using liquid chromatography–mass spectrometry analysis (J. S. Lee et al., 2016). A study by Jang et al (2014) reported a *de novo* heterozygous deletion in 2q33.3–q34 in a 15-months-old female with RTT-like features. This region is considered to harbour genes associated with autism and neurodevelopmental delay. The identified deletion using aCGH was 5.9Mb in size (206,048,173–211,980,867) and contained 34 genes. The candidate genes identified were *NRP2, ADAM23, KLF7, CREB1, MAP2, UNC80,* and *LANCL1* (Jang, Chae, & Kim, 2015). Baasch et al revealed a novel *de novo* missense variant in the *SCN2A* gene [c.5645G>T; p.(Arg1882Leu)] in a 5-year old girl with heterogeneous disorder including epilepsy and severe ID using WES (Baasch et al., 2014). In 2014, a *de novo* variant in another gene associated with ASD, *TBL1XR1* [c.209G>A; p.(Gly70Asp)] was detected in a Japanese girl diagnosed with West Syndrome, RTT-like and autistic features by Saitsu et al using WES. This change was within an F-box-like motif at a residue which is evolutionarily conserved (Saitsu et al., 2014).

Moreover, investigation of 68 Chinese children with unexplained early infantile epileptic encephalopathy (EIEE) (including West syndrome, Dravet syndrome, Ohtahara syndrome and unclassified EIEE) found *de novo* pathogenic variants in *CDKL5* (n = 2), *STXBP1* (n = 2), *SCN1A* (n = 3), *KCNQ2* (n = 2) and *SCN8A* (n = 1). They found that *SCN1A* (38.5%) was the most recurrent gene with *de novo* pathogenic variants followed by *STXBP1* (15.4%), *CDKL5* (15.4%), *KCNQ2* (15.4%), *ARX* (7.7%) and *SCN8A* (7.7%) respectively (Arafat et al., 2017). Archer et al (2006) investigated the possible involvement of the *NTNG1* gene in a large cohort of 115 *MECP2* mutation-negative RTT individuals following the identification of a child with atypical RTT who had epileptic seizures and a chromosomal re-arrangement disrupting this gene. In total, four sequence variants were identified. However, three of them were intronic with no effect on splicing and one was synonymous variant. Thus, all the identified variants in *NTNG1* gene were unlikely pathogenic (Archer et al., 2006).

1.6 Key molecular pathways dysregulated in NDDs

The molecular characterization of the key proteins behind the complex phenotype observed in individuals with NDDs is critical to understanding the complete patho-physiology behind these disorders. This understanding would further open up avenues for potential targeted drug discoveries to possibly correct the perturbed molecular pathways. Perturbed gene expression regulation is the common pathway mainly causing various NDDs including RTT.

1.6.1 Chromatin remodelling and NDD

Chromatin, a DNA-histone complex composed of repeating nucleosome units, is a highly dynamic structure providing structural organisation of the genome and regulating gene expression; thereby maintaining cellular integrity at specific developmental stages. The plasticity in chromatin structure is affected by a) DNA methylation, b) histone modifications, c) ATP-dependent chromatin regulation (as discussed below). These processes are regulated by a group of specialised remodelling enzymes known as chromatin remodelers (Tyagi, Imam, Verma, & Patel, 2016).

DNA methylation, the major form of DNA modification, is carried out by DNA methyltransferases (DNMTs) that transfer the methyl groups at CpG islands in DNA (Mirabella, Foster, & Bartke, 2016). Traditionally, the link between chromatin regulation and RTT was evident through the global regulatory role of MeCP2 that binds to the methylated CpG dinucleotides for transcriptional repression and activation, chromatin remodelling and RNA splicing (Sanfeliu, Kaufmann, Gill, Guasoni, & Tropea, 2019). With regards to HDAC-dependent transcriptional repression by MeCP2, this results in selective binding of the MBD domain of MeCP2 to the CpG islands, which recruits the chromatin remodelling complexes comprising of transcriptional co-repressor Sin3A, Brahma (BRM; a SWI/SNF-related chromatin remodelling protein) and HDACs to the TRD domain (Figure 1.7; A). Collectively this complex de-acetylates lysine residues on histone tails, thereby resulting in the chromatin structure to further condense thereby restricting the availability of promoter regions for transcription of the downstream genes (Della Ragione, Vacca, Fioriniello, Pepe, & D'Esposito, 2016). In addition, MeCP2 activates the regional transcription of downstream genes in specific brain regions of *Mecp2* mouse models. Studies have also showed that up to 85% of genes were found to be downregulated in the hypothalamus of *Mecp2* null and upregulated in *Mecp2* duplication mice, further supporting the role of Mecp2 as a transcriptional activator (Chahrour et al., 2008). Additionally, the promoter of activated genes was found to be associated with Mecp2 and cAMP response element-binding protein 1 (CREB-1) (Figure 1.7; B) (Chahrour et al., 2008).

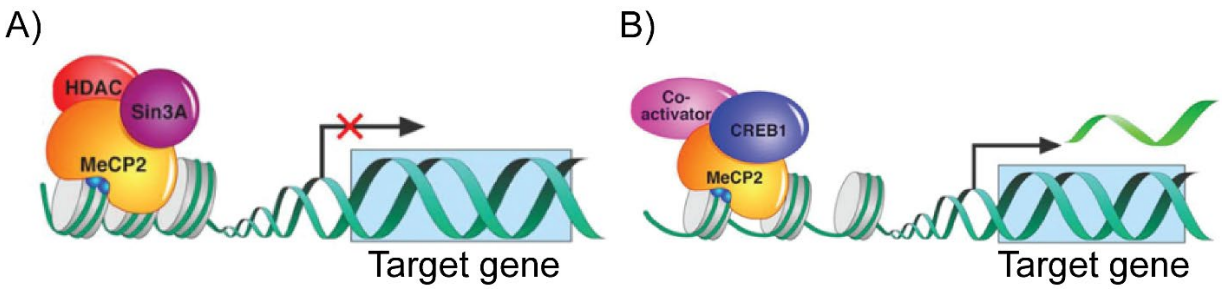


Figure 1.7 MeCP2 as transcriptional repressor and activator.

A) Binding of MeCP2 recruits the transcriptional co-repressor complex consisting of Sin3A, HDAC and Brahma (not shown), that de-acetylates lysine residues on histone tails and thus repressing downstream gene expression due to condensed conformation of chromatin. B) MeCP2 acting as transcriptional activator in the hypothalamus by associating with its direct target CREB1. Image taken from (Chahrour et al., 2008).

Histone modifications involve the reversible acetylation and/or methylation of histone tails that protrude from the nucleosome to control DNA transcription, replication and repair. The positively charged lysine residues within the histone tails in nucleosomes are acetylated and deacetylated during chromatin remodelling and gene regulation. When the acetyl group is added to histones by histone transferases (HATs), the positive charge of lysine residue is neutralized and reduces the interaction of histones with the negatively charged DNA backbone, causing DNA to unwind. Histone deacetylases (HDACs) perform the opposite function by removing the acetyl group from histone tails, tightly packing the chromatin (heterochromatin), and further preventing the access of transcription factors for gene expression (Figure 1.8). Interestingly, several HATs can acteylate non-histone proteins as well (Portela & Esteller, 2010).

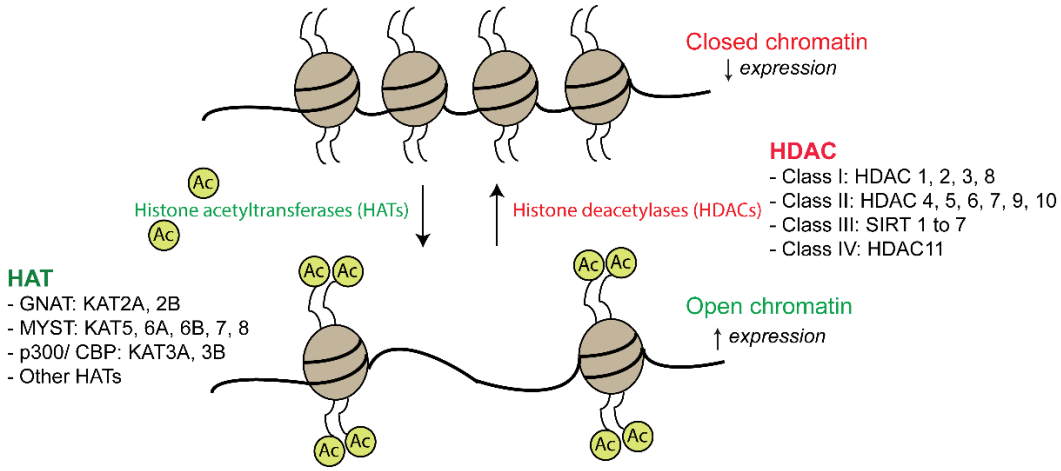


Figure 1.8: The role of histone deacetylases and histone acetyltransferases in chromatin remodelling.

Addition of acetyl groups by histone acetyltransferases (HATs) results in open chromatin structure that results in increased gene expression whereas, histone deacetylases (HDACs) remove the acetyl group from the lysine residues on histone tails, thereby increasing interactions between histones and DNA, resulting in closed chromatin conformation and reduced gene expression. The key proteins of HATs and HDAC families are also shown. Ac = acetyl group.

In mammals, 18 HDAC enzymes are further categorized into two families (Classical HDACs and Sirtuins) and four different classes: Class I consisting of HDAC1, HDAC2, HDAC3, and HDAC8 (Rpd3-like proteins), Class II containing HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10 (Hda1-like proteins), Class III comprising of SIRT1, SIRT2, SIRT3,

SIRT4, SIRT5, SIRT6, and SIRT7 (Sir2-like proteins) and finally Class IV with a single member, HDAC11 (Seto & Yoshida, 2014). All known HATs are divided into three main families: the Gcn5-related acetyltransferase (GNAT) family, the MYST family (MOZ, Ybf2/Sas3, Sas2, and Tip60) and the p300/CBP family. There also exists a loosely defined category of "other HATs" that mainly consists of acetylating proteins that differ considerably in their structure to other HAT family proteins. In human, the HATs have been classified as lysine (K) acetyltransferases (KATs) and thus an alternating nomenclature exists in the reported literature as shown in (Table 1.3) (Wapenaar & Dekker, 2016).

Table 1.3: Members of different HAT families and alternative nomenclature.

Protein	Alternative nomenclature	HAT family	Associated disease
KAT2A	Gcn5	GNAT	-
KAT2B	PCAF		-
KAT5	TIP60	MYST	-
KAT6A	MOZ, MYST3		Mental retardation, autosomal dominant 32
KAT6B	MORF, MYST4		Genitopatellar syndrome Say-Barber-Biescker-Young-Simpson syndrome
KAT7	HBO1, MYST2		-
KAT8	MOF, MYST1		-
KAT3B	p300	p300/CBP	Rubinstein-Taybi syndrome
KAT3A	CBP		Rubinstein-Taybi syndrome (Miller & Rubinstein, 1995)
KAT4	TAF1, TBP	Other HATs (Transcriptional co-activators)	X-linked Dystonia-Parkinsonism
KAT12	TIFIIIC90		-
KAT13A	SRC1	Other HATs (Steroid receptor co-activators)	-
KAT13B	SCR3, AIB1, ACTR		-
KAT13C	p600		-
KAT13D	CLOCK		-
KAT1	HAT1	Other HATs (Cytoplasmic)	-
HAT4	NAA60		-

Adapted from (Wapenaar & Dekker, 2016).

ATP-dependent chromatin remodelling: The third pathway in the epigenetic gene expression regulation involves **ATP-dependent chromatin remodelling** complexes that are distinct from the DNA and histone modifying enzymes. They have the ability to modify chromatin structure using energy through ATP hydrolysis to disrupt the strong interactions between DNA and histones, thereby opening nucleosome structure for access to transcription factors (Hota &

Bruneau, 2016). Along with the ATPase domain, these protein have non-catalytic subunits that regulate the binding and activity of these complexes. According to the ATPase subunit, the chromatin-remodelling complexes are usually classified into four main families; switch/sucrose-non-fermenting (SWI/SNF), imitation switch (ISWI), chromodomain-helicase-DNA binding (CHD), and inositol requiring 80 (INO80) families. All the remodelers share a conserved ATPase domain belonging to superfamily 2 (SF2) of DEAD/H-box helicases that is split into two tandem RecA-like folds (DEXX and HELIC) (Figure 1.9) (Byrd & Raney, 2012). Additional domains characteristic of a particular family are involved in recruiting the remodelling complex to chromatin, nucleosome interaction or ATPase activity regulation (Tyagi et al., 2016).

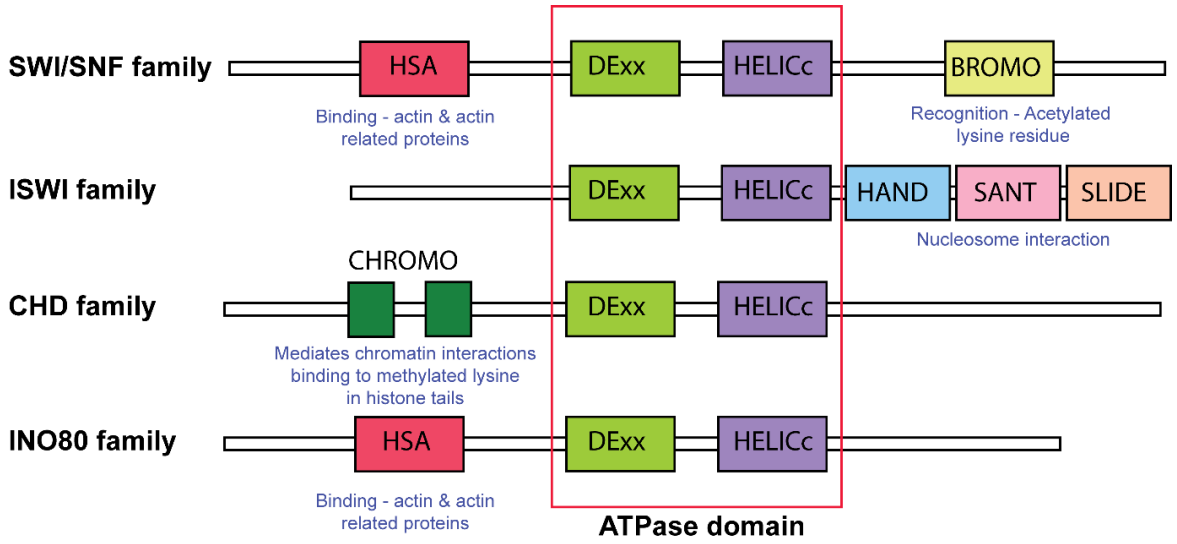


Figure 1.9 Chromatin remodelers family of proteins.

Schematic showing the domain structure of four families of ATP-dependent chromatin remodelers. The ATPase domain (DExx and HELICc) is well-conserved among the four families. BROMO domain is only present in SWI/SNF family, whereas the CHROMO domain is specific for the CHD family. The HAND, SANT and SLIDE domains in particular are present in only ISWI family. Image adapted from (Tyagi et al., 2016).

Owing to the substantial importance of epigenetic modifications for gene expression regulation, it is not surprising that any slight perturbations in these developmentally regulatory pathways could contribute to a variety of disorders including NDDs. Pathogenic variants in a few HAT family members have been found to be implicated in various disorders of cognitive

dysfunction (Sheikh, 2014). Heterozygous variants in *CBP/ KAT3A* and *p300/KAT3B* have been reported in individuals with Rubinstein-Taybi syndrome (RTS) (Petrij et al., 1995; Roelfsema et al., 2005). The prevalence of RTS is 1:125,000 live births and is characterized with severe intellectual disability, cardiac defects, dysmorphic facial features, prominent broad thumbs and large toes, and increased susceptibility to cancers originating from neural cells (Miller & Rubinstein, 1995; Roelfsema & Peters, 2007). Various *in vivo* studies have highlighted the critical role of p300 and CBP in gene expression activation during differentiation into mature neuronal cell lineages (Viosca, Lopez-Atalaya, Olivares, Eckner, & Barco, 2010; Wang et al., 2010). Seventy five percent of $p300^{+/-}$ mice showed impaired growth and craniofacial abnormalities, traits that are often seen in RTS individuals (Viosca et al., 2010). Reduced binding of CBP at the promoters of the neuronal and glial genes was found in cortical cultures from E11.5-E12.5 of $Cbp^{+/-}$ mice, and decreased histone acetylation (H3K9/14) was observed, which may be relevant to the intellectual disability phenotype observed in RTS individuals (Wang et al., 2010). Similarly, pathogenic variants in other HAT family members, Lysine (K) acetyltransferase 6A (KAT6A/ MOZ/ MYST3) and its paralogue Lysine (K) acetyltransferase 6B (KAT6B/ MORF/ MYSR4), have been found to be associated with syndromic developmental delay and/or intellectual disability (Kennedy et al., 2019; Kim et al., 2017; Lundsgaard et al., 2017). *De novo* heterozygous variants in *KAT6A* have been previously identified in individuals with *KAT6A*-related intellectual disability and alterations in *KAT6B* have been known to cause Say-Barber-Biesecker-Young-Simpson syndrome characterized by intellectual disability and congenital anomalies (Clayton-Smith et al., 2011) and Genitopatellar Syndrome, which is a rare disorder associated with intellectual disability, genital anomalies and craniofacial defects (Campeau et al., 2012; Simpson et al., 2012). Similarly, chromatin remodelling proteins may be dysfunctional in NDD and epileptic disorders. Pathogenic variants in ATP-dependent CHD family members, *CHD1*, *CHD2*, *CHD4*, *CHD6*, *CHD7* and *CHD8*, have been reported in individuals with a range of neurological conditions including development delay, ASD, ID and epilepsy (Schoberleitner et al., 2019). In addition, MeCP2 have been shown to interact with α-thalassemia mental retardation X-linked protein (ATRX) which is a member of SWI/SNF family and is critical for heterochromatin formation (De La Fuente, Baumann, & Viveiros, 2011; Della Ragione, Filosa, Scalabri, & D'Esposito, 2012). Pathogenic variants in *ATRX* result in X-linked syndromic and

non-syndromic intellectual disability with varying degree of gonadal dysgenesis, further reinforcing the association between RTT, NDD and chromatin dysregulation.

Thus, various chromatin regulators have been found to be implicated in NDDs. However, the exact role of HATs and various ATP dependent chromatin regulators still needs to be further elucidated in RTT. However, owing to significant phenotypic overlap between individuals with RTT and those with NDD caused by genetic defects in these chromatin regulators, as well as the global regulatory role of MeCP2, it is possible that further studies may uncover missing links between these epigenetic regulatory pathways in RTT.

1.6.2 Microtubule associated protein defects in NDDs

Microtubules are an essential component of the cytoskeleton. They are hollow but rigid rod-like filaments that have a highly dynamic structure. They form crucial networks that play a variety of roles, including determination of the shape of a cell to provide a scaffold for coordinated transport of organelles and vesicles, as well as separation and organisation of chromosomes during cell division (Glotzer, 2009; Stanhope & Ross, 2015). The monomer unit is tubulin which is a heterodimer consisting of α-tubulin and β-tubulin bound together by guanosine triphosphate (GTP). These tubulin dimers polymerise to form protofilaments, and thirteen protofilaments arrange parallel to each other around a hollow core (Figure 1.10). Due to the specific orientation of the tubulin dimer, microtubules have directional polarity, with a fast growing plus end and a slow growing minus end. This polarity determines the direction of movement of cargo along microtubules (Cooper, 2000). The hydrolysis of GTP bound to β-tubulin results in depolymerisation of the microtubule. There is a dynamic behaviour, during which the GTP-bound tubulin dimer is added on the plus end and guanosine diphosphate (GDP)-bound tubulin is removed from the minus end. This dynamic phenomenon is known as “treadmilling” (Cooper, 2000).

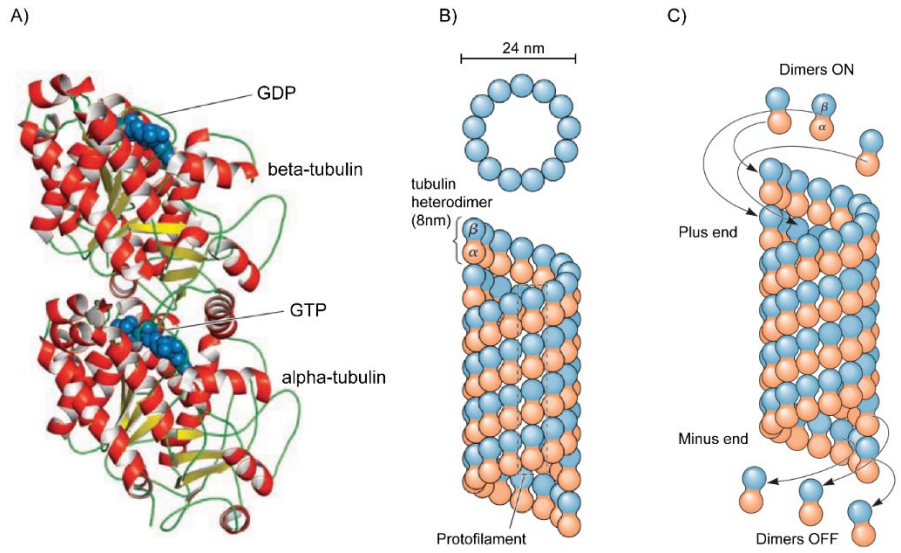


Figure 1.10: Structure of αβ-heterodimer of tubulin and model of GTP dependent tread milling process.

A) αβ-heterodimer of tubulin (PDB: 1JFF). B) Protofilament consists of alternating αβ-heterodimers and microtubules consists of 13 parallel protofilaments forming a 24nm wide hollow structure. C) "treadmilling" phenomenon in microtubules where a GTP-bound tubulin dimer is added on the plus end, and GDP-bound tubulin is removed from the minus end (Garrett & Grisham, 2005).

The microtubule cytoskeleton plays a critical role during neuronal development, particularly in neurite formation and axon specification, axon elongation and branching, dendritic arborisation and synapse formation and intracellular transport via motor proteins for neuronal function and maintenance (Lasser, Tiber, & Lowery, 2018).

Kinesin superfamily

Variants in the kinesin superfamily (KIF) genes, encoding ATP-dependent molecular motor proteins critical for brain function, are increasingly being linked with intellectual disability-related NDDs (Lasser et al., 2018). In mammals, there are over 45 KIF genes that are divided into 15 KIF families, termed KIF1 to KIF14B, based on phylogenetic analysis (Figure 1.11). These families are further broadly divided into three types (N-, M- and C-) of kinesins based upon the position of the well conserved motor domain that is critical for ATP hydrolysis and movement of KIF proteins (Hirokawa, Noda, Tanaka, & Niwa, 2009). Pathogenic variants in *KIF4A* and *KIF5C* were found in individuals with intellectual disability-related NDD in two

unrelated families, and these variants were shown to interrupt the synaptic function in primary rat hippocampal neurons (Willemsen et al., 2014). *De novo* likely pathogenic variants in another kinesin family member 1A (*KIF1A*) gene have been found in individuals with intellectual disability and various other NDD specific clinical features, including cognitive dysfunction, spastic paraplegia and neurodegeneration. This complex phenotype is commonly referred to as KIF1A-associated neurological disorder (KAND) (L. Boyle & W. Chung, 2017; Cheon et al., 2017; Citterio et al., 2015; Erlich et al., 2011; Hotchkiss et al., 2016; Klebe et al., 2012; Krenn et al., 2017; J. R. Lee et al., 2015; Ohba et al., 2015; Okamoto et al., 2014; Wang, Zhang, et al., 2019; Yoshikawa et al., 2019).

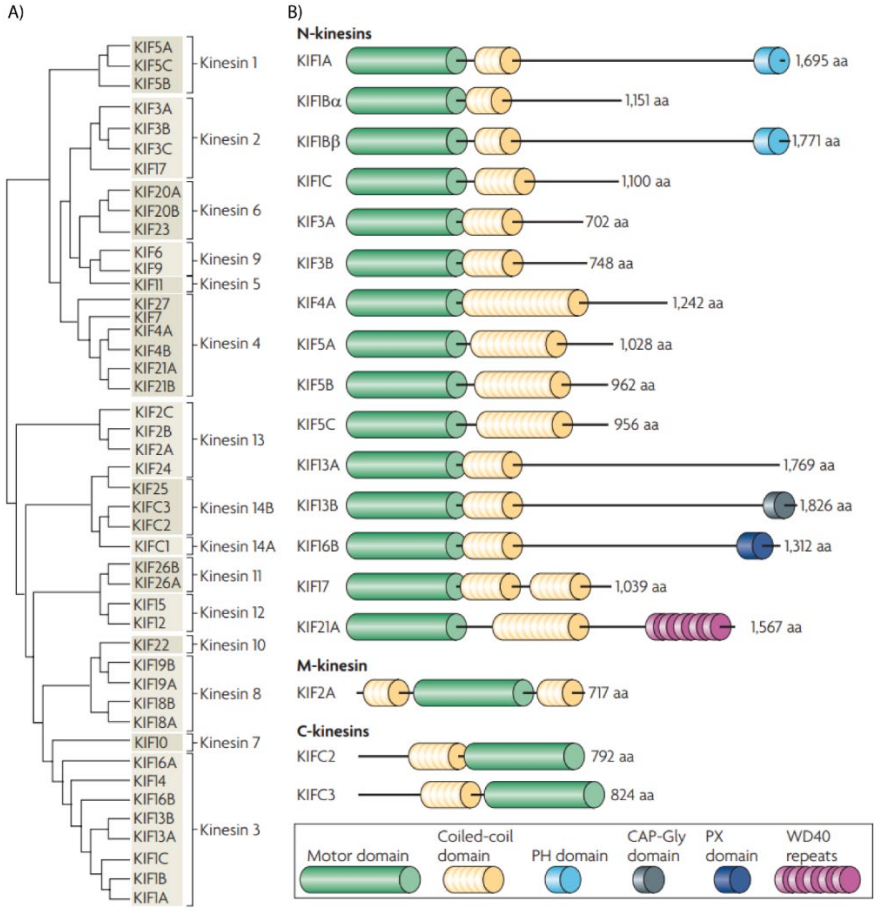


Figure 1.11: Kinesin superfamily (KIFs) A) genes in the mouse genome B) their domain structure.

A) Phylogenetic tree showing 45 KIF proteins divided into 15 families (KIF1 to KIF14B). B) Domain structure of N-, M- and C- kinesins showing key domains in KIF. The critical motor domain (green) binds to the microtubules and move step-wise using ATP to transport specific cargos attached at the PH-domain (Cyan); from (Hirokawa et al., 2009).

KIF1A is an N-kinesin, since the motor domain (amino acids 1-361) is located at the N-terminus of the protein and consists of ATP hydrolysis and microtubule binding domains (Figure 1.12). Downstream of the motor domain lie other functional domains critical for dimerization and kinesin activation including neck coil (NC) followed by coiled coil segment CC1, a fork-head domain (FHA) and another two coiled coil segments CC2 and CC3. The C-terminal pleckstrin domain (PH) binds with the cargo to transport it from cell body towards end of the neurite terminals (Figure 1.12) (Hirokawa et al., 2009). The motor domain of KIF1A has been found to be highly conserved across various species.

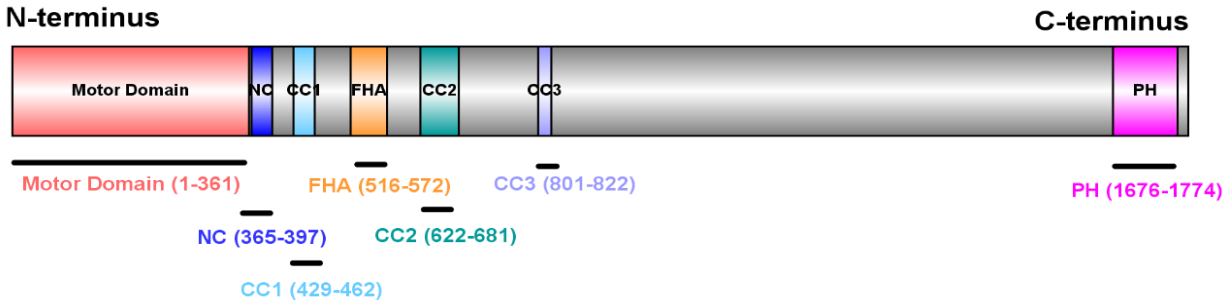


Figure 1.12: Domain structure of full length kinesin family member 1A (KIF1A).

Motor domain (1-361), NC: Neck coil (365-397), CC1: Coiled coil 1 (429-462), FHA: fork-head domain (516-572), CC2: Coiled coil 2 (622-681), CC3: Coiled coil 3 (801-822), PH: pleckstrin domain (1676-1774). (NP_001230937.1). Image generated using NCBI coordinates in Illustrator for Biological Sequences (IBS) program (W. Liu et al., 2015).

KIF1A is a neuron-specific microtubule plus end–directed motor protein that is involved in fast anterograde transport of specific synaptic vesicle precursors along the axons such as Brain Derived Neurotrophic Factor (BDNF), synaptophysin, synaptotagmin and RAB3A. Moreover, KIF1A has been shown to directly interact with the liprin-α, which further associates KIF1A with other pre-synaptic protein cargos such as GRIP, GIT1, RIM and βPIX (Okada, Yamazaki, Sekine-Aizawa, & Hirokawa, 1995; Shin et al., 2003). Hence, KIF1A plays a crucial role in the maintenance and proper functioning of mature neurons, and reduced KIF1A levels in the brain can disturb critical neuronal transport further. *KIF1A* is a disease-associated gene in the spectrum of neurological disorders, commonly known as KAND, and may be implicated in the pathogenesis of RTT (L. Boyle & W. Chung, 2017; Wang, Zhang, et al., 2019).

1.6.3 Perturbed microtubule dynamics in RTT

Many studies have provided evidence that microtubule stability and trafficking is often dysregulated in neuronal cells with variations in the RTT associated genes *MECP2* and *CDKL5*. (Baltussen et al., 2018; Barbiero et al., 2017; Chahrour et al., 2008; Delepine et al., 2016; Delepine, Nectoux, Bahi-Buisson, Chelly, & Bienvenu, 2013; Gold, Lacina, Cantrill, & Christodoulou, 2015; Jordan, Li, Kwan, & Francke, 2007; Kaufmann, Naidu, & Budden, 1995; Maezawa, Swanberg, Harvey, LaSalle, & Jin, 2009; Nectoux et al., 2012; Williams et al., 2014; Xu, Kozikowski, & Pozzo-Miller, 2014). In 2006, Pelka et al found significantly decreased expression of kinesin family member 1B (Kif1b) in the hippocampus of *Mecp2*^{tm1Tam} mice (Table 1.4) (Pelka et al., 2006). Kif1b is involved in microtubule-based trafficking of essential cargoes such as mitochondria from the cell body towards the nerve terminals. Furthermore, Gibson et al (2010) found a key microtubule associated protein to be dysregulated in the cerebral cortex of RTT individuals (Gibson et al., 2010), namely Dynamin I (*DNM1*). It is involved in trafficking of vesicles along microtubules and was found to be upregulated two-fold in the frontal cortex of RTT individuals whereas levels were significantly reduced in occipital cortex when compared with controls. In 2012, Nectoux et al identified a significant down-regulation of stathmin-like 2 (*STMN2*; also named *SCG10*), in fibroblasts derived from classic RTT individual as well as in the cerebellum of Mecp2 deficient mice, which is involved in microtubule assembly as well as dynamics. Furthermore, they found that the reduced levels of *STMN2* were only evident in astrocytes and not neurons (Nectoux et al., 2012). Interestingly, the level of *STMN2* was also found to be reduced in hypothalamus of Mecp2-deficient *Mecp2*$^{tm1.1Bird}$ mice (Chahrour et al., 2008) and cerebellum of *Mecp2*$^{tm1.1Jae}$ and *Mecp2*$^{tm1.1Bird}$ mice (Table 1.4) (Jordan et al., 2007).

MeCP2 directly modulates the expression of KIF1A cargo protein, BDNF, in an activity-dependent manner (Zhou et al., 2006). A significant reduction in BDNF mRNA and protein has been reported in Mecp2-deficient mice and RTT individuals (W. Li, Calfa, Larimore, & Pozzo-Miller, 2012). Furthermore, severe RTT-like symptoms and enhanced seizure risk were observed in mutation-positive RTT individuals with single nucleotide polymorphism (SNP) in the *BDNF* gene (Ben Zeev et al., 2009). Moreover, Xu et al (2014) also found impaired dendritic BDNF trafficking along microtubules in cultured hippocampal neurons of *Mecp2*

knockout mice. Recently a *de novo* heterozygous truncating variant [NM_001244008.1: c.275_276insAA; NP_001230937.1: p.(Cys92*)] in the motor domain of KIF1A has been also reported in a single female case presenting with classic RTT; however no functional validation supported this finding (Wang, Zhang, et al., 2019). This cumulative evidence suggests a crucial association between BDNF and KIF1A defects, MeCP2 function and RTT pathophysiology.

The role of MeCP2 in microtubule dynamics was first uncovered by Delepine et al (2013) where they observed faster microtubule depolymerisation and reduced acetylation under cold-induced stress in fibroblasts from RTT individuals with pathogenic variants p.(Ala2Val), p.(Arg294*) and p.(Arg168*), thereby highlighting the involvement of MeCP2 in microtubule stability (Delepine et al., 2013). HDAC6 reduces microtubule stability by removing crucial acetyl groups from microtubules and a selective HDAC6 inhibitor, Tubastatin A has been shown to increase the levels of acetylated tubulin and interestingly to increase the velocity of the BDNF vesicles with no neuro-toxic effects (Xu et al., 2014). Gold et al (2015) further investigated the functional aspects of microtubules and levels of HDAC6 in fibroblasts from two male RTT individuals [NM_004992.3: c.806delG; p.(Gly269Alafs) and c.419C>T; p.(Ala140Val)] and in a *Mecp2* knock-in mouse model (*Mecp2*^{T158A}). Reduced levels of acetylated tubulin were observed in both individual fibroblasts and mouse cortical neurons. Increased HDAC6 levels were reported in cortical neurons from the Mecp2-deficient mice. Moreover, the level of tubulin acetylation was rescued using Tubastatin A with no increase in MeCP2 expression (Gold et al., 2015). Along with neurons, astrocytes also play a crucial role in the pathophysiology of RTT (Maezawa et al., 2009; Williams et al., 2014). Recently, Delepine et al (2016) reported alterations in the microtubule-based vesicle transport in Mecp2-deficient mouse astrocytes as well as from MeCP2 p.(Arg294*) iPSC-derived astrocytes. Moreover, they found another natural microtubule stabilising agent, Epothilone D restored *in vitro* microtubule stability in Mecp2/ MeCP2-deficient mouse and human iPSC derived astrocytes (Delepine et al., 2016). All these recent studies further reinforce the role of Mecp2/MeCP2 in microtubule dynamics and cargo trafficking in neurons.

Table 1.4 Key microtubule associated genes dysregulated with MeCP2 deficiency

Gene	Gene Name	Function	Sample	Sample type	↑/↓	References
MAP2	Microtubule associated protein 2	Stabilizes microtubule	RTT individual	Neocortex	↓	(Kaufmann et al., 1995)
Kif1b	Kinesin family member 1B	Vesicle transport along microtubules	*Mecp2*^{tm1Tam} mice	Hippocampus	↓	(Pelka et al., 2006)
DNM1	Dynamin 1	Vesicle transport along microtubules; Hydrolysis of GTP; Synthesis of microtubule bundles	RTT individual	frontal cortex	↑	(Gibson et al., 2010)
STMN2 (also SCG10)	Stathmin-like 2	Most abundant in growth cones during development microtubule-destabilizing activity	RTT individual Mecp2-deficient mice	Fibroblasts Cerebellum (astrocytes)	↓ -	(Nectoux et al., 2012)
Dync1h1	Dynein cytoplasmic 1 heavy chain 1	Vesicle transport along microtubules	Mecp2-deficient mice *Mecp2*$^{tm1-1Bird}$	Rostral brain regions	↓	(Roux et al., 2012)
Dctn1	Dynactin 1	Binds to both microtubules and cytoplasmic dynein	Mecp2-deficient mice *Mecp2*$^{tm1-1Bird}$	Rostral brain regions	↓	(Roux et al., 2012)

- : No change

Several studies have shown that CDKL5 deficiency has a negative impact on various cellular functions, such as axon specification and outgrowth, dendritic branching as well as spine structure, which are all particularly influenced by microtubule dynamics (Barbiero et al., 2017)). CDKL5 has been found to interact with Shootin1, a key neuronal polarization regulator, to regulate axon formation in the growing neurites. Silencing of CDKL5 reduced Shootin1 phosphorylation in hippocampal neurons (Nawaz et al., 2016; Toriyama et al., 2006). Moreover, reduced CDKL5 levels resulted in decreased binding of the cytoplasmic linker protein 170 (CLIP170) with microtubules in cells and in the growth cones of *Cdkl5*-KO

neurons, thus further affecting the axon specification and outgrowth (Barbiero et al., 2020). CDKL5 also regulates dendritic arborisation through Rac1 which is an actin remodeller, and it is interesting to note that dendritic branching defects are often observed in individuals with CDKL5 deficiency disorder (Barbiero et al., 2017; Chen et al., 2010; Swiech et al., 2011). CDKL5 deficiency resulted in increased binding of microtubule associated protein 1S (MAP1S), with the dendritic microtubules further resulting in an increased duration of EB3 at the microtubule ends, thus resulting in less dynamic microtubule plus ends, thereby affecting microtubule dynamics (Baltussen et al., 2018; Ramkumar, Jong, & Ori-McKenney, 2018). Functionally, CDKL5 has been shown to regulate the structure and function of highly dynamic dendritic spines that rely on the remodelling of actin cytoskeleton structure, and are critical for excitatory synaptic transmission (Barbiero, De Rosa, & Kilstrup-Nielsen, 2019). In the absence of CDKL5, microtubules were found to be less dynamic, and failed to invade into the dendritic spines, and thus impacting the activity dependent formation of mushroom-shaped spines, which correlated with reduced synaptic strength (Zhu & Xiong, 2019).

Therefore various microtubule-associated proteins have been found to be impacted in RTT as a consequence of the deficiency in MeCP2 or CDKL5. This suggests that there could be a further repertoire of microtubule associated proteins perturbed by variants in RTT-associated genes still to be uncovered.

1.7 Current research: Hypothesis and Aims

RTT is one of the major NDDs and despite recent advances in NGS techniques, a significant fraction of atypical RTT/RTT-like individuals do not have a genetic diagnosis, preventing accurate genetic counselling. This not only has a huge impact on the families of affected individuals but, in the absence of a clear understanding of the underlying molecular basis of their disorder, the development of targeted therapies is prevented.

Our research team has a long-standing commitment to RTT research, and in 2015 using WGS, we identified a novel *de novo* heterozygous missense variant [NM_001244008.1: c.744C>A; p.(Asp248Glu)] in the motor domain of KIF1A in one classic RTT female. Variations in *KIF1A*

were traditionally linked to several neurological disorders, collectively known as KAND, but had not been linked to RTT.

Hypothesis:

Genetic variations in kinesins and other novel genes contribute to the pathophysiology of RTT and other related neurodevelopment disorders.

To test this hypothesis, three main aims were outlined:

Aim 1: To study the evolving role of kinesinopathies in the pathogenesis of RTT.

We screened 42 out of 88 collected genetically undiagnosed RTT/RTT-like individuals using a range of NGS techniques including WES, WGS and with a custom designed NGS targeted sequencing panel. The sequencing panel consisted of 136 genes including all the coding regions of all known kinesin proteins as well as genes known to be associated with RTT, Angelman syndrome and intellectual disability related NDDs. The pathogenicity of identified *KIF1A* variants was further evaluated using a range of functional assays.

Aim 2: Evaluation of the functional implications and phenotypic manifestations of *KIF1A* variants associated with KAND.

This aim focused on functional analysis of likely pathogenic *de novo* heterozygous missense variants in *KIF1A* identified in a cohort of KAND individuals collected from a number of international collaborators. All individuals were carefully reviewed according to Neul's diagnostic criteria for RTT (Neul et al., 2010) in order to identify whether any of these might meet a classification of RTT. We characterised the functional consequences of *KIF1A* variants using two main functional assays; neurite tip accumulation assay and *in-vitro* microtubule gliding assay. These methods are widely used to determine the functional consequences of

individual specific variants on the mobility of KIF1A protein on synthetic microtubules and in cultured cells.

Aim 3: Next generation sequencing of other mutation-negative RTT individuals.

This aim focused on an in-depth analysis of genes that we identified in mutation-negative RTT/RTT-like individuals using NGS technologies in Aim 1. We categorised likely pathogenic genes into three main categories:

a) Known disease genes previously associated with RTT
b) Known disease genes not previously associated with RTT
c) Novel candidate disease genes not previously reported with any disorder

Functional studies were undertaken for a range of these genes, to further evaluate the pathogenicity of identified variants.

Overall, this PhD project aimed to confirm whether kinesin family member genes could be contributing to the phenotype in RTT/RTT-like individuals. In addition, using genomic and functional genomic techniques, we have uncovered other NDD-related genes in mutation-negative RTT individuals, which has improved the overall genetic diagnostic rate for RTT individuals.

Chapter 2

Chapter 2 Materials and Methods

2.1 Ethics, consents and permissions

All procedures used in this study were approved by the human research ethics committees of the participating institutes with written consents obtained from the legal guardians of each participant and in accordance with the Helsinki Declaration of 1975, as revised in 2000.

2.2 Subjects and genetic analysis

The genetically undiagnosed individuals were recruited in collaboration with various institutes as described in respective chapters. Three different NGS technologies (WGS, WES or targeted panel sequencing) were used to identify the likely causative gene in these individuals. NGS was performed using genomic DNA isolated from blood (unless otherwise specified), and as detailed in the respective chapters, either at the Translational Genomics Unit, Victorian Clinical Genetic Services (VCGS) after assessing the quality of DNA using Qubit and TapeStation 2200 (Agilent Technologies, Cat No. G2965AA), or at the participating institutes as discussed in respective chapters.

Whole Genome Sequencing (WGS)

WGS libraries were prepared using Nextera™ DNA Flex Library Prep and loaded onto a HiSeq 4000 sequencer (Illumina; HiSeq control Software v3.4.0) and 2 x 150 bp paired-end sequencing was performed at VCGS. Samples passed sequencing QC with > =80% bases with at least Q30, and mean coverage of at least 27-fold coverage.

Whole Exome Sequencing (WES)

WES SureSelectQXT Clinical Research Exome libraries were prepared and loaded onto a NextSeq 500 sequencer (Illumina; NextSeq control Software v2.1.031) and 2 x 150 bp paired-

end sequencing was performed. Samples passed sequencing QC with > 89.7% bases with at least Q30, and mean coverage of at least 100-fold.

Targeted panel sequencing

SureSelect custom DNA Panel (Agilent Technologies) with 50,439 probes for 136 genes (Appendix A, List A2) was designed using Agilent SureDesign software based on the GRCh37/hg19 Human Genome assembly. The targeted gene list contained known RTT associated genes, all KIF family member genes and genes implicated in other NDDs. The panel consisted of the exonic region of all 136 genes along with an additional 10 bp into the intron at 5' and 3' ends of the exons to identify potential splice site variants. The libraries were prepared using SureSelect QXT Target Enrichment System (Agilent technologies) as per the manufacturer's instructions. Sequencing was performed on an Illumina MiSeq using 150bp paired-end reads. Fifty nanograms of the individual DNA at a minimum concentration of 25ng/µl was used for targeted sequencing at VCGS.

2.2.1 Variant filtering

For the NGS sequencing performed at VCGS, Illumina's bcl2fastq2 converter (v2.17.1.14) was used to convert the raw sequencing data into FASTQ format and was further processed using Cpipe (http://cpipeline.org/) (Sadedin et al., 2015). Alignment to the reference genome (GRCh37) was performed in order to generate annotated variant calls within the target region (coding exons +/- 2bp). Identified variants were annotated against all gene transcripts as per the HGNC recommended transcript according to HGVS nomenclature and uploaded onto Leiden Open (source) Variation Database (LOVD) for further analysis (Fokkema et al., 2011). Variants with low quality (<100), high genomic mean allelic frequency (>5%), intronic variants, untranslated region (UTR) variants, synonymous and low impact variants were filtered using an "unknown singleton inheritance" filter in LOVD. First, any RTT associated genes were inspected to find any previously missed known RTT variants or novel variants in known RTT associated genes (Christodoulou, Grimm, Maher, & Bennetts, 2003; Krishnaraj et al., 2017). Variants in known intellectual disability associated genes were further filtered using a pre-curated phenotype-specific gene list by the VCGS (Appendix A, List A3). A total of 1,064 genes implicated in syndromic as well as non-syndromic intellectual disability were

covered in this gene list. The remaining variants were further prioritised based upon the inheritance pattern and impact of the variant in the order of; early stop gain, frameshift and missense variants.

2.2.2 *In silico* analysis and variant classification

The population frequency of the identified variants was determined using the Genome Aggregation Database (gnomAD) (v2.1.1; http://gnomad.broadinstitute.org/) (Karczewski et al., 2019). Pathogenicity of the variants was assessed using various *in silico* tools including MutationTaster (v2.0; www.mutationtaster.org/) (Schwarz, Cooper, Schuelke, & Seelow, 2014), PolyPhen-2 [v2.2.2r398; http://genetics.bwh.harvard.edu/pph2/; score range: 0.0 (tolerated) to 1.0 (deleterious)] (Adzhubei et al., 2010) and Sorting Intolerant From Tolerant [SIFT; v6.2.1; http://sift.jcvi.org/; score range: 0.0 (deleterious) to 1.0 (benign); prediction cut-off = 0.05] (Sim et al., 2012). The classification scheme reported by The American College of Medical Genetics and Genomics (ACMG) was used to classify the variants as pathogenic, likely pathogenic or benign (Richards et al., 2015). Grantham score (range 0 to 215) was assigned to the amino acid substitutions, and the a missense change was considered significant when the score was greater than 100 (Grantham, 1974). In addition, the disease-specific allele frequency of the identified variant was evaluated using ClinVar, the public archive of human genetic variants (https://www.ncbi.nlm.nih.gov/clinvar/) (Landrum et al., 2018).

2.2.3 Structural modelling

The *in silico* three dimensional (3D) structural modelling was performed on the canonical transcript using HOPE (URL: http://www.cmbi.ru.nl/hope/) (Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010). In some cases, the position of the residue affected was shown in 3D structure images generated using PyMOL (http://www.pymol.org/) molecular visualization software (Schrödinger, 2010).

2.3 Molecular biology techniques

2.3.1 DNA extraction

Genomic DNA was extracted from cultured fibroblasts (<10 passages) using QIAamp DNA Mini Kit (Qiagen; Cat. no. 51304) kit as per the manufacturer's protocol for cultured cells, eluted in ultra-pure water and quantified using and stored at -30ºC until use.

2.3.2 Constructs and mutagenesis

The sequences of all the constructs used in this project were analyzed using SnapGene Viewer software (version 4.0.2; https://www.snapgene.com/snapgene-viewer/; GSL Biotech).

KIF1A studies

A constitutively-active dimeric KIF1A motor (amino acids 1-393) containing the motor domain (amino acid 1-361), neck linker, and neck coil domains and dimerized via a GCN4 leucine zipper (LZ) was a kind gift from Prof Kristen Verhey (University of Michigan, USA) (Hammond et al., 2009). For the neurite tip accumulation assay, KIF1A(1-393)-LZ was tagged with three tandem mCitrine (mCit) fluorescent proteins at the C-terminus [KIF1A(1-393)-3xmCit] (Appendix B, Figure B1). For the microtubule gliding assay fluorescent mNeonGreen (mNG; Allele Biotech) and AviTag (amino acid sequence GLNDIFEAQKIEWHE) were attached at the C-terminus [KIF1A(1-393)-LZ-mNG-AviTag] (Appendix B, Figure B2). The HA-tag labelled bacterial biotin ligase BirA was co-transfected with KIF1A(1-393)-LZ-mNG-AviTag into COS-7 cells to biotinylate the expression motor domain of the KIF1A protein at the AviTag site (Yue et al., 2018). The biotin tag was utilised for the immobilization of KIF1A(1-393) protein from COS-7 whole cell lysates on the surface of NeutrAvidin for microtubule gliding assays.

KIF23 (or MKLP1) studies

Full length KIF23 (also called MKLP1) construct tagged with pmCherry at the C-terminus (pmCherry-C1-MKLP1) (Appendix B, Figure B3), which was a gift from Masanori Mishima

(Addgene plasmid#70154; http://n2t.net/addgene:70154; RRID:Addgene_70154), was used for the functional analysis of *KIF23* variant (Joseph, Hutterer, Poser, & Mishima, 2012).

Site-directed mutagenesis using QuikChange lightning mutagenesis kit (Agilent technologies; Cat No. 210519) was used to introduce individual-specific variants into the KIF1A and KIF23 constructs according to the manufacturer's instructions.

2.3.3 Bacterial cells and chemical competent cell preparation

Lysogeny broth (LB) medium, routinely supplied by institutional core facilities (Murdoch Children's Research Institute; MCRI) was used to maintain bacterial cells. The media was solidified using agar-agar powder (15 g/L) for LB-agar plates. For selecting bacterial cells with plasmid constructs, the media was supplemented with appropriate antibiotics with concentrations as: 50 µg/ml kanamycin, 100 µg/ml ampicillin. Frozen stock of *Escherichia coli* (*E. coli*) TOP10 and the protocol for making competent cells was based on Rubidium Chloride method adapted from Hanahan et. al. and was kindly shared by Dr. Paul Kalitsis (MCRI) (Hanahan, Jessee, & Bloom, 1991). Aliquots of chemically competent bacterial cells were aliquoted, flash frozen in liquid nitrogen and stored at -80ºC until further use.

2.3.4 Transformation of chemical competent *E. coli* cells

Chemically competent *E. coli* cells were thawed on ice prior to adding 100ng of plasmid DNA. The DNA-cell mix was incubated on ice for 30 mins before subjecting to heat shock at 42ºC for 30 secs, followed by incubation on ice for 5 mins. Approximately 450µl of pre-warmed LB-medium (without any antibiotic) was then added to the transformed cells, followed by incubation at 750rpm at 37ºC for 1 hour. The recovered cells were then spread on the pre-warmed LB-agar plates with appropriate antibiotics and incubated at 37ºC overnight.

2.3.5 Plasmid DNA preparation and assessment

Following transformation and overnight incubation, a single colony of the *E. coli* was inoculated into 75 ml of LB-media with appropriate antibiotics and shaken overnight at 37ºC.

The next day, cells from the overnight culture were harvested by centrifuging at 8000g for 20 mins and plasmid DNA was isolated using QIAGEN Plasmid Midi Kit (Qiagen; Cat No. 12145) following the manufacturer's instructions. The concentration and quality (A260/280) of the prepared plasmid was determined using NanoDrop 2000c UV-Vis spectrophotometer (Thermo Fisher Scientific) and the correct size was assessed using agarose gel electrophoresis.

2.3.6 Agarose gel electrophoresis

Agarose gels were prepared by adding 1% (wt/vol) agarose powder (Sigma; Cat no. A9539) in 1X Tris-borate-EDTA (TBE) buffer routinely supplied by MCRI core facilities. Agarose was then dissolved into 1x TBE by heating prior to adding SYBR safe DNA staining dye (10,000X; Thermo Fisher Scientific; Cat no. S33102) and cast in the BioRad agarose gel electrophoresis trays. DNA loading buffer (New England Biolabs Inc.; Cat. no. B7025S) was added to the DNA samples and samples were run at 100V for approximately 20-30 mins. The pre-stained 100 bp (Thermo Fisher Scientific; Cat. no. SM0243) or 1 kb DNA ladders (Thermo Fisher Scientific; Cat. no. SM0313) were also loaded as molecular weight standard markers.

2.3.7 Mammalian cell lines and transfections

All the mammalian cell lines were grown according to the standard protocols established by the MCRI Tissue Culture Service laboratory.

Skin fibroblasts and COS-7 (African green monkey kidney fibroblast, American Type Culture Collection) cell lines were grown in DMEM media (Dulbecco's Modified Eagle's Medium (Sigma; Cat. no. D6546; with 3.7g/L $NaHCO_3$, 4.5g/L D- glucose and 0.584g/L L-glutamine) media supplemented with 10% (vol/vol) fetal bovine serum (FBS; Gibco) and 100 units/mL penicillin, and 100 µg/mL streptomycin at 37°C in a humidified incubator and 5% CO_2. For passaging, cells were washed with warm sterile 1X PBS and covered with enough 0.025% trypsin/EDTA (supplied by MCRI) solution prior to incubating them at 37ºC for 3 mins or until the cells detached from the surface after examining under an inverted microscope. To inhibit further activity of trypsin, media containing c10% FBS was added, usually three

times the volume of trypsin used. Cells were then harvested by centrifugation at 300g for 5 mins and resuspended with fresh media before plating.

SH-SY5Y (human neuroblastoma) and Neuro2A (mouse neuroblastoma) cell-lines were grown in RPMI growth media (Sigma; Cat. no. R0883; with 3.7g/L $NaHCO_3$, 2.0g/L glucose and 0.3g/L L-glutamine) supplemented with 10% (vol/vol) FBS, 1X Pen-Strep at 37 °C in a humidified incubator and 5% CO_2. When cells were approximately 80% confluency, cells were passaged using 0.25% trypsin (supplied by MCRI). Cells were differentiated using RPMI differentiation media (RPMI supplemented with 1% FBS, 1X Pen-Strep and 10µM retinoic acid (RA; Sigma; R2625)). The differentiation media was refreshed every 24 hours for four days.

After careful optimisation of various commercial transfection reagents, cells were transfected with the appropriate construct using Lipofectamine 2000 (ThermoFisher Scientific; Cat No. 11668019) following the manufacturer's instructions for 24 hour. Detailed information regarding plating density, amount of DNA and Lipofectamine 2000 is discussed in respective chapters.

2.3.8 Polymerase Chain Reaction (PCR)

Genomic DNA fragment amplification

Primers for genomic DNA amplification were designed using a web-based tool, Primer-3 (http://bioinfo.ut.ee/primer3-0.4.0/) using standard parameters including primer size [18 to 27 nucleotide (nt)], primer melting temperature (T_m; 57ºC to 63ºC), maximum T_m difference (= 1ºC), CG clamp (= 1), GC content (40 - 60%), maximum self-complementary score (= 8) and maximum 3' self-complementary score (= 3) (Untergasser et al., 2012). De-salted primers at 100µM concentration were ordered (Macrogen, Seoul) and diluted to 10µM before using in the PCR reactions. The primer sequences used are listed in Appendix A (Table A1).

Standard DNA fragment amplification reactions were carried out in 25 µl reactions consisting of template DNA (100 ng) along with Hot Start Taq DNA polymerase (Qiagen; Cat. no. 203203; 1U), forward/ reverse primers (0.5µM), deoxyribonucleotide triphosphates (dNTPs; 0.2mM each), magnesium chloride ($MgCl_2$; 1.5 mM) and PCR buffer (1X). The PCR reaction was run in a Bio-Rad thermal cycler at 95°C for 15 mins (1 cycle) for initial denaturation followed by 35 cycles at 95°C for 30 secs, 60°C for 30 secs and at 72°C for 1 min. The final extension was carried out at 72°C for 10 mins.

Site-directed mutagenesis

Primers flanking individual specific variation site were designed using web-based QuikChange Primer Design Program available online (www.agilent.com/genomics/qcpd) as per the software instructions. The primer sequences are listed in Appendix A (Table A1).

Fifty microliters of the PCR reaction was prepared consisting of dsDNA template (100ng), reaction buffer (1X), forward/ reverse primers (125ng each), dNTP mix (1 µl), QuikSolution reagent (1.5 µl), QuikChange Lightning Enzyme (1 µl) using QuikChange Lightening kit (Agilent; Cat. no. 210518). The reaction was run at 95°C for 2 mins (1 cycle) for initial denaturation, followed by 18 cycles at 95°C for 20 secs, 60°C for 10 secs and at 68°C for varying times dependent on the size of the construct. Final extension was carried out at 68°C for 5 mins. The time of extension was calculated as per the size of construct, with a 5-kb construct requiring 2.5 mins per cycle at 68ºC; KIF1A(1-393)-3xmCit (5.913 kb: ~3 mins), KIF1A(1-393)-LZ-mNG-AviTag (6.049 kb; ~3 mins).

Colony Screening following mutagenesis

After bacterial transformation step (as listed in Section 2.3.4), at least ten single colonies were picked from the LB-agar plate and grown overnight in LB media with appropriate antibiotics. Twenty microliters of overnight culture was then added to 100 µl of ultrapure water and boiled at 95°C for 10 mins. The lysate was then centrifuged at 13,000g for 2 mins. Ten microliters of supernatant was used for the PCR in a standard genomic DNA fragment amplification reaction,

as discussed above. PCR products were cleaned up and sent for Sanger sequencing to confirm mutagenesis.

PCR reactions were further optimised by adjusting the annealing temperature when the standard conditions were not successful. The PCR products were then run on 1% (w/v) agarose gel as described in Section 2.3.6.

2.3.9 Gel purification and PCR product clean-up and gel

The PCR reactions were cleaned up for sequencing to remove unused dNTPs and primers that interfere with the downstream applications. Five microliters of PCR product was mixed with 2 µl of ExoSAP-IT™ (ThermoFisher Scientific; Cat no. 78201.1.ML), incubated at 37^0C for 30 mins followed by 80^0C for 15 mins in the thermocycler. Following clean-up, samples were submitted for sequencing.

On the occasions where multiple separated DNA fragments on 1% agarose gel were observed after gel electrophoresis, despite careful PCR optimisation, gel purification of the band of required size was carried out. Briefly, the appropriate band was identified based on size and sliced out using a clean blade under ultraviolet light using the Dark Reader DR-45M transilluminator (Clare Chemical Research). Gel extraction of the DNA fragments was then carried out using QIAquick Gel Extraction Kit (Qiagen; Cat No. 28704) following the manufacturer's instructions. Further clean-up was not carried out before submitting for Sanger sequencing.

2.3.10 Sanger sequencing

Both the sense and anti-sense strands were sequenced by submitting 15µl of reaction containing 1 µl of purified PCR product (or 1 µg of construct), and 1 µl of either forward or reverse primer (10µM) in ultra-pure water to the Sequencing Service and Development Platform at VCGS (MCRI) for Sanger sequencing. All the ABI sequencing files were analyzed using CodonCode Aligner software (demo version v. 8.0.2; https://www.codoncode.com/aligner/; CodonCode Corporation).

2.4 Protein studies

2.4.1 Extraction and quantification of protein from fibroblasts

The whole cell extracts were prepared by culturing control and individual fibroblasts to confluency, rinsing twice with ice-cold PBS prior to extracting in cold RIPA buffer (10 mM Tris-Cl (pH 8.0), 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl, 1mM PMSF (phenylmethylsulfonyl fluoride) and 1x protease inhibitor cocktail (Roche)) with short gentle sonication on ice before incubating on ice for 30 mins. The whole cell extract was collected as a supernatant after centrifuging at 18,000g for 20 mins at 4°C, then stored at -80°C. Following the manufacturer's instructions, the Pierce BCA Protein Assay Kit (Thermo Scientific; Cat No. 23225) was used to determine the protein concentration of the lysates. The absorbance at 562 nm was measured using a FLUOstar Optima Microplate Reader (BMG Labtech).

2.4.2 SDS- polyacrylamide gel electrophoresis (PAGE)

Extracted protein lysates (5-25 µg) were dissolved in 3X loading buffer [65mM Tris-HCl-pH 6.8, 26% glycerol, 2.1% SDS, 0.01% bromophenol blue] supplemented with β-mercaptoethanol (in the ratio of 500:60) and heated at 95ºC for 5 mins followed by centrifugation at 10,000g for 1 min. The prepared samples were then loaded onto SDS-PAGE reducing gels (10-12% or gradient 4-15% acrylamide gels based on the protein being analyzed) along with Precision Plus Protein™ Dual Color Standard (BioRad; Cat No. 161-0374). The separation of the proteins was carried out under denaturating conditions in running buffer (25mM Tris-Cl, 192mM glycine, 0.1% SDS) at a constant voltage of 110V for 45-60 mins. Separated proteins were then transferred to methanol activated Hybond PVDF (0.45µm) membranes (GE Healthcare; Cat no. 10600023) using pre-chilled transfer buffer (25mM Tris-Cl, 192mM glycine, 20% methanol) at 100V for 1 hour at 4ºC. Non-specific sites on the membrane were blocked using blocking solution made up of either using 5% (w/v) skim milk or 5% bovine serum albumin (BSA) in 1X TBS [50 mM Tris-HCl (pH 7.6), 150 mM NaCl] supplemented with 0.01% of Tween-20 (TBST) for 30 mins at room temperature. Membranes were then incubated overnight at 40ºC with appropriate primary antibodies (Table 2.1). The following day, the membrane was washed five times with 1X TBST for 5 mins each at room

temperature prior to blocking with blocking solution for 30 mins at room temperature. Appropriate secondary anti-mouse or anti-rabbit horseradish peroxidase-conjugated antibody (GE Healthcare) at dilution 1:10,000 was then added to the membrane for 1 hour at room temperature followed by five washes for 5 mins each with TBST solution. Membranes were then developed using ECL Primer western blotting detection reagent (GE Healthcare; Cat. no. RPN2232) and exposed to autoradiography film (GE Healthcare, Cat. no. 28906837). The developed autoradiography films were scanned into .tiff files using CanoScan 9000F Mark II (version 1.1.2; Canon Inc.) software and band densities were quantified using Image J software (Abramoff, Magelhaes, & Ram, 2004).

Table 2.1 List of antibodies used in this thesis

Protein detected	Primary antibody	Species	Conditions	Epitope	Theoretical molecular weight
Chromodomain helicase DNA-binding protein 8 (**CHD8**)	CHD8 (D3C1)^ (Cell Signalling Technologies; Cat no. 11891)	Rabbit	1:2000 Overnight 4ºC	C-terminal (unknown sequence)	290 kDa
Nucleoporin-188 (**NUP188**)	NUP188 antibody ^^ (Bethyl Laboratories Inc; Cat no. A302-322A-M)	Rabbit	1:1000 Overnight 4ºC	C-terminal (amino acids 1600-1650 of human NUP188; NP_056169.1)	188 kDa
SET Nuclear Proto-Oncogene (**SET**)	Anti-SET/TAF-I antibody (Abcam; Cat. no. 1183)	Rabbit	1:500 Overnight 4ºC	N-terminal (amino acids 3-18 of human SET; NP_001116293.1)	~37 kDa
SET Nuclear Proto-Oncogene (**SET**)	Anti-SET antibody (GeneTex; Cat. no. GTX113834)	Rabbit	1:500 Overnight 4ºC	Centre of human SET; NP_001116293.1, unknown sequence	33 kDa Observed: 42/44 kDa
Glyceraldehyde-3-phosphate dehydrogenase (**GAPDH**)	Anti-GAPDH Sigma Aldrich; (Cat no. G9545)	Rabbit	1:10,000 1 hour RT	C-terminal (amino acids 314-333 of mouse GAPDH NP_001276655.1)	~36 kDa
β-Actin	β-Actin (C4) (Santa Cruz Biotechnology; Cat. no. sc-47778)	Mouse	1:1000 1 hour RT	Unknown	~43 kDa

^: This antibody cross-reacts with a protein of unknown origin at 140 kDa; ^^: This antibody cross-reacts with a protein of unknown origin at 100 kDa; RT = room temperature

2.4.3 *In-vitro* neurite tip accumulation assay

The optimised density of SH-SY5Y cells ($5*10^4$ cells per well) was plated on sterile glass coverslips coated with 3.3% rat collagen (Corning; Cat. no.354236) and 2% matrigel (Corning; Cat. no. 354248) (recipe optimised by Dr. Wendy Gold) in RPMI media without serum and antibiotics in a 12 well plate. After 24 hours, cells were transfected with the appropriate construct using Lipofectamine 2000 (Life Technologies; Cat. no. 11668019) according to the manufacturer's instructions. Neuronal differentiation of the transfected cells was then induced using RA (as mentioned in Section 2.3.7). Cells were then washed with warm 1X phosphate buffer saline (PBS; 137mM NaCl, 5.4mM KCl, 1.28mM NaH_2PO_4, 7mM Na_2HPO_4; pH 7.4), fixed in 4% (vol/vol) paraformaldehyde (PFA; Sigma; Cat. no. 158127) in PBS for 20 mins at room temperature before mounting on glass slides using Prolong Gold with DAPI (Thermo Fisher Scientific; Cat. no. P36935). Images were acquired using Zeiss AxioVision or LSM780 fluorescence microscope with 40X objective (zoom factor =2). ImageJ was used to quantify the mean florescence intensity (MFI) of the expressed fluorescent protein along the neurite length and cell body of cells. Statistical analysis on the average MFI was performed using Mann Whitney in Prism software (GraphPad). All experiments were performed in triplicate and repeated three times prior to data analysis.

2.4.4 *In-vitro* microtubule gliding assay

COS-7 cells were plated at an optimised density of $1*10^5$ per well in a 6-well plate and co-transfected with optimised 1 µg of KIF1A(1-393)-LZ-mNG-AviTag (Figure 2.1) and an equal amount of 1 µg of HA-BirA construct (optimised by our collaborators; Prof Kristen Verhey) using 5 µl of Lipofectamine 2000.

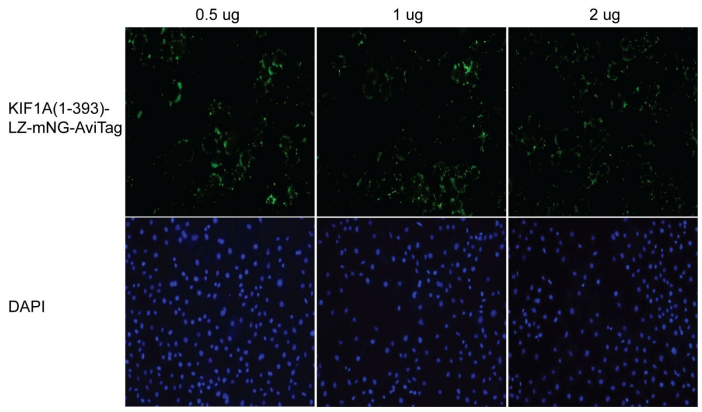


Figure 2.1: Optimisation of the transfection amount of the KIF1A(1-393)-LZ-mNG-AviTag in COS-7 cells for the in-vitro microtubule gliding assay.

Total of 1x105 cells per well were plated on glass coverslips in 6 well plates. 0.5µg, 1µg and 2µg of wild type KIF1A(1-393)-LZ-mNG-AviTag construct was transfected into COS-7 cells. After 16 hours, cells were washed, fixed with ice cold methanol for 5 minutes at room temperature and imaged using AxioVision Zeiss microscope.

Confirmation of biotinylation of KIF1A protein:

Cells were washed three times with warm 1XPBS after 16 hours of transfection, fixed using 4% PFA (vol/vol) for 20 mins at room temperature. After fixation, cells were again washed three times with 1X PBS and permeablised with pre-chilled methanol at -30ºC for 5 mins. After washing again three times with 1X PBS, cells were blocked

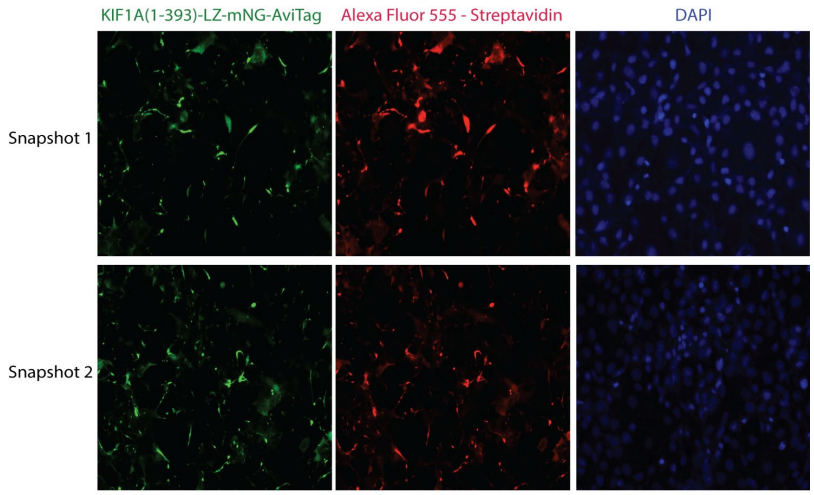


Figure 2.2: Biotinylation confirmation of expressed KIF1A protein.

COS-7 cells were co-transfected with 1 µg each of KIF1A(1-393)-LZ-mNG-AviTag and HA-BirA. After 16 hours, probed with Alexa Fluor® 555 streptavidin. Green = KIF1A, red = biotinylated KIF1A and blue = DAPI.

overnight with 1% casein (Sigma; Cat no. C7078; prepared in 1X PBS) at 4ºC. The following day cells were incubated with Alexa Fluor® 555 streptavidin (1:3000; Thermo Fisher Scientific; Cat. no. S21381; gifted by Wicks Laboratory, Walter and Eliza Hall Institute of Medical Research) for 1 hour at room temperature. After washing with 1X PBS, cells were mounted on a glass slide using Prolong Gold with DAPI and imaged using a Zeiss AxioVision microscope. The biotinylation was confirmed by determining coincident staining of Alexa Fluor® 555 streptavidin (red) with fluorescent KIF1A(1-393) (green) (Figure 2.2).

Preparation of COS-7 lysate with biotinylated KIF1A(1-393) protein: After 16 hours of transfection (optimised using KIF1A(1-393)_6X His tag; data not shown), COS-7 cells were washed with 1X PBS, trypsinized and harvested by centrifuging at 3000g at 4ºC. After washing the cell pellet with 1X PBS, ice-cold lysis buffer (30 µl/ well of cultured cells; 25 mM HEPES/KOH, pH 7.4, 115 mM potassium acetate, 5 mM sodium acetate, 5 mM $MgCl_2$, 0.5 mM EGTA, and 1% [vol/vol] Triton X-100) supplemented with 1 mM ATP, 1 mM PMSF (Sigma; 10837091001), and protease inhibitor cocktail (Thermo Fisher Scientific; A32965) was added. The resuspended cells were incubated at 4ºC for 30 mins. The soluble fraction was collected after centrifugation (20,000g) for 15 mins at 4ºC, snap frozen in single use 30 µl aliquots, and stored at -80ºC. The stored lysates remained active for at least 8 weeks.

Polymerisation of microtubules: The lyophilized unlabelled tubulin from porcine brain (Cytoskeleton Inc; Cat no. T240-A) and rhodamine-labelled tubulin (Cytoskeleton Inc; Cat no. TL590M-A) were reconstituted to a final concentration of 10 mg/ml using ice-cold general tubulin buffer (80mM PIPES, 2mM $MgCl_2$, 2.5mM EGTA pH 7.0) supplemented with fresh 1mM guanosine triphosphate (GTP; Cytoskeleton; Cat. no. BST06), snap frozen in single use aliquots, and stored at -80ºC. Unlabelled tubulin and rhodamine-labelled tubulin were mixed together in the molar ratio of 10:1, respectively in general tubulin buffer supplemented with 1mM GTP and 5% glycerol prior to incubating the mix at 37ºC for 30 mins to induce tubulin polymerization. Paclitaxel (Cytoskeleton Inc; Cat. no. TXD01) was then added at increasing concentrations (2µM, 20 µM and 200 µM) with 10 mins of incubation at each step in order to maintain the polymerised state of the fluorescent microtubules. The synthesised microtubules were separated from unpolymerised monomers using 40% glycerol cushion with

ultracentrifugation at 300,000 g for 10 mins at room temperature. The pelleted microtubules were re-suspended in warm general tubulin buffer with 1mM GTP, 5% glycerol and 20µM paclitaxel. Microtubule concentration (in mg/ml) was measured using the Nanodrop at OD 280nm and was recorded as 10mg/ml. The value was converted into µM using equation 1 mg/ml tubulin = 10 µM; considering the approximate molecular weight of αβ-tubulin heterodimer as 100,000 Daltons. Hence, the concentration of the prepared stock of microtubules was 100µM. Polymerization was confirmed by imaging under AxioVision Zeiss microscope (Figure 2.4) and 10 µl of microtubules was stored at room temperature and remaining were snap-frozen as single use aliquots and stored at -80°C.

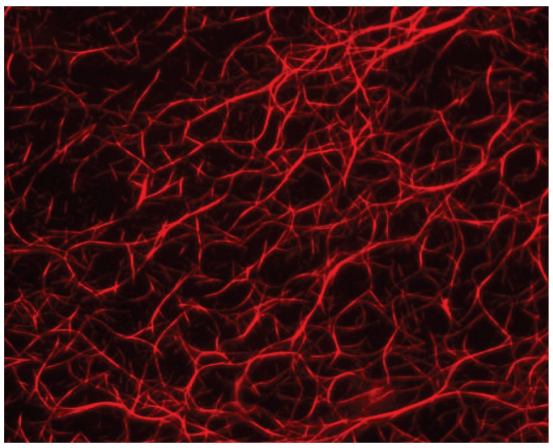


Figure 2.3: Microtubule polymerisation confirmation

Rhodamine labelled microtubules (red) were polymerised using unlabelled and labelled tubulin monomers (10:1). Image at 40X was captured using AxioVision Zeiss microscope.

Flow chamber preparation and gliding assay: A 10 µl flow chamber on a glass slide was prepared by attaching two strips of double-sided tape spaced 5 mm apart and inverting a clean #1.5 coverslip (Thermo Fisher Scientific) over the top. The biotinylated KIF1A proteins in the prepared lysate were immobilized on the attached coverslip by sequential addition of 100 µl (unless specified) of a series of different solutions, each incubated for 5 mins with coverslip touching the base of the surface as shown in Figure 2.4 [Adapted from (Yue et al., 2018)]. First, 1 mg/ml BSA-biotin (Sigma; Cat no. A8549) was added into the flow

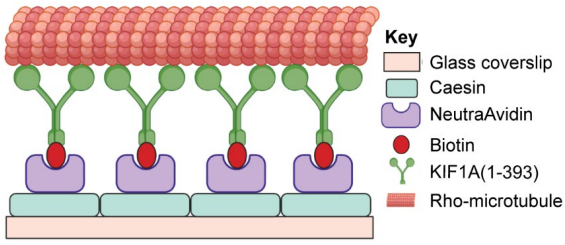


Figure 2.4: Flow chamber set up for *in-vitro* microtubule gliding assay

Sequential addition of chemicals in the order of casein, NeutraAvidin, Biotinylated KIF1A(1-393) & rhodamine (Rho) labelled microtubules were added to the glass flow chamber prepared by inverting clean glass coverslip on two parallel double sided tape mounted on a glass slide at a distance of 5mm.

chamber, followed by blocking solution containing 0.5mg/ml casein in BRB80 buffer (80mM PIPES, 1mM $MgCl_2$ and 1mM EGTA) supplemented with 10µM paclitaxel. Next, NeutrAvidin (0.5 mg/ml; Thermo Fisher Scientific; Cat. no. 31000) was added to the flow chamber followed by addition of blocking solution before finally adding 30 µl of COS-7 cell lysates expressing KIF1A motor protein, supplemented with 2mM ATP to maintain the activity of KIF1A protein.

Fresh motility mix containing 0.5µl of prepared taxol-stabilized rhodamine-labelled microtubules, 2mM ATP, 2mM $MgCl_2$ and enzymatic oxygen scavenging system (10mM glucose, 200µg/ml glucose oxidase, 1mM dithiothreitol, 100 µg/ml catalase), to avoid oxidation-mediated microtubule degradation due to oxidation, was prepared just prior to imaging. Then 10µl of the motility mix was added into the chamber and clear varnish was applied to seal the ends of the flow cell. The movement of fluorescent microtubules was captured using a Leica DMIB6000 total internal reflection microscopy (TIRF) microscope fitted with a 100X U PlaApo 1.49 NA oil immersion objective and solid state 488, 561 nm laser line at room temperature. Signals were collected using EMCCD camera (Andor iXon 897 Ultra) and sequential images were acquired continuously at 50 msec per frame for 4000 frames. The distance versus time graphs (Kymographs) using a width of 3 pixels were then generated by using the multi-kymograph plugin in ImageJ (Abramoff et al., 2004). The movement of individual microtubules was tracked by manually drawing lines on the maximum intensity projections of the sequential images as previously published (Yue et al., 2018). The average velocity (µm/sec) of moving microtubule was calculated by taking the ratio of distance (x-axis) and time (y-axis). This experiment was repeated by using cell lysates prepared in three independent batches on separate days. The data was analysed using ANOVA and Mann-Whitney tests in Prism software (GraphPad) and the data was presented as mean ± standard deviation (SD).

Chapter 3

Human Mutation
Variation, Informatics, and Disease
OFFICIAL JOURNAL HGV HUMAN GENOME VARIATION SOCIETY

Phenotypic spectrum of *de novo* missense variants in kinesin family member 1A (*KIF1A*): expansion to include Rett syndrome and Rett-like patients

Journal:	*Human Mutation*
Manuscript ID	humu-2019-0594.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Kaur, Simranpreet; Murdoch Children's Research Institute, Royal Children's Hospital, Brain and Mitochondrial Research Group; University of Melbourne, Department of Paediatrics Van Bergen, Nicole; Murdoch Childrens Research Institute, Brain and Mitochondrial Research Group; University of Melbourne, Department of Paediatrics and Child Health Verhey, Kristen; University of Michigan Medical School Department of Cell and Developmental Biology Nowell, Cameron; Monash Institute of Pharmaceutical Sciences. Monash University, Drug Discover Biology Budaitis, Breane; University of Michigan, Cellular and Molecular Biology Program Yang, Yue; University of Michigan Medical School Department of Cell and Developmental Biology Massey, Sean; Murdoch Childrens Research Institute, Brain and Mitochondrial Research Group Ellaway, Carolyn; Children's Hospital at Westmead, Western Sydney Genetics Program; University of Sydney, Disciplines of Genetic Medicine and Child and Adolescent Health, Sydney Medical School Brunetti-Pierri, Nicola; Federico II University of Naples, Pediatrics; Telethon Institute of Genetics and Medicine, Cappuccio, Gerarda ; Federico II University of Naples, Pediatrics; Telethon Institute of Genetics and Medicine, Bruno, Irene; University of Campania Luigi Vanvitelli, Department of Precision Medicine Boyle, Lia; Columbia University Irving Medical Center, Division of Molecular Genetics Nigro, Vincenzo; University of Campania Luigi Vanvitelli, Department of Precision Medicine Torella, Annalaura; University of Campania Luigi Vanvitelli, Department of Precision Medicine Roscioli, Tony; Sydney Children's Hospital Randwick, Centre for Clinical Genetics; University of New South Wales - Randwick Campus, Prince of Wales Clinical School; University of New South Wales - Randwick Campus, Neuroscience Research Australia Cowley, Mark; Garvan Institute of Medical Research, Kinghorn Centre for Clinical Genomics; University of New South Wales, St Vincent's Clinical School; University of New South Wales, Children's Cancer Institute, Lowy

	Cancer Research Centre Sonawane, Rhea; Deakin University, Faculty of Science, Engineering and Built Environment Burton, Matthew; Murdoch Children's Research Institute, Flow Cytometry and Imaging facility Schönewolf-Greulich, Bitten; Copenhagen University Hospital, Genetic Counselling Clinic Kennedy Center Tumer, Zeynep; Copenhagen University Hospital, Applied Human Molecular Genetics, Kennedy Center, Department of Clinical Genetics Chung, Wendy; Columbia University Medical Center, Departments of Pediatrics and Medicine Gold, Wendy; Kids Research Institute, Molecular Neurobiology Research Laboratory; The University of Sydney, School of Medical Sciences and Discipline of Child and Adolescent Health, Faculty of Medicine and Health; Children's Hospital at Westmead, Kids Neuroscience Centre, Kids Research Christodoulou, John; Murdoch Childrens Research Institute, Brain and Mitochondrial Research Group; University of Melbourne, Department of Paediatrics and Child Health
Key Words:	Rett syndrome, MECP2, KIF1A, neurite tip accumulation, microtubule, kinesin

Phenotypic spectrum of *de novo* missense variants in kinesin family member 1A (*KIF1A*): expansion to include Rett syndrome and Rett-like patients

Simranpreet Kaur[1,2*], Nicole J. Van Bergen[1,2*], Kristen J. Verhey[3], Cameron J. Nowell[4], Breane Budaitis[5], Yang Yue[3], Carolyn Ellaway[6,7], Nicola Brunetti-Pierri[8,9], Gerarda Cappuccio[8,9], Irene Bruno[10], Lia Boyle[11], Vincenzo Nigro[10], Annalaura Torella[10], Tony Roscioli[12,13], Mark J. Cowley[14,15,16], Sean Massey[1], Rhea Sonawane[17], Matthew D. Burton[18], Bitten Schonewolf-Greulich[19], Zeynep Tümer[19], Wendy Chung[20], Wendy A. Gold[21,22,23^], John Christodoulou[1,2,6,24^]

1. Brain and Mitochondrial Research Group, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Australia
2. Department of Paediatrics, University of Melbourne, Melbourne, Australia
3. Department of Cell and Developmental Biology, University of Michigan Medical School, Ann Arbor, MI, USA
4. Drug Discover Biology, Monash Institute of Pharmaceutical Sciences. Monash University, VIC, Australia
5. Cellular and Molecular Biology Program, University of Michigan Medical School, Ann Arbor, MI, USA
6. Discipline of Genomic Medicine, , School of Medical Sciences, Faculty of Medicine and Health, University of Sydney, NSW, Australia
7. Western Sydney Genetics Program, Children's Hospital at Westmead, Westmead, NSW, Australia
8. Department of Translational Medicine, University of Naples "Federico II", Italy
9. Telethon Institute of Genetics and Medicine, Pozzuoli (NA), Italy
10. Department of Precision Medicine, University of Campania "Luigi Vanvitelli", Naples, Italy
11. Division of Molecular Genetics, Columbia University Irving Medical Center, NY, USA
12. New South Wales Health Pathology, Randwick, Sydney, Australia
13. Neuroscience Research Australia, University of New South Wales, Sydney, Australia
14. Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Sydney, Australia
15. St Vincent's Clinical School, UNSW Sydney, Sydney, Australia
16. Children's Cancer Institute, Lowy Cancer Research Centre, UNSW, Sydney, Australia

17. Faculty of Science, Engineering and Built Environment, Deakin University, Melbourne, Australia
18. Flow Cytometry and Imaging facility, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Australia
19. Kennedy Center, Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Glostrup, Denmark
20. Departments of Paediatrics and Medicine, Columbia University Medical Center, NY, USA
21. Molecular Neurobiology Research Laboratory, Kids Research, Children's Hospital at Westmead, Westmead, NSW, Australia
22. Kids Neuroscience Centre, Kids Research, Children's Hospital at Westmead, Westmead, NSW, Australia
23. School of Medical Sciences and Discipline of Child and Adolescent Health, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia
24. Victorian Clinical Genetics Services, Royal Children's Hospital, VIC, Australia

* = co-first authors, ^ = co-last authors

Correspondence

John Christodoulou

Brain and Mitochondrial Research Group, Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Victoria 3052 Australia

Email: john.christodoulou@mcri.edu.au

ABSTRACT

Rett Syndrome (RTT) is a severe X-linked neurodevelopmental disorder caused mainly by pathogenic variants in the methyl-CpG-binding (*MECP2*) gene; however, up to 40% of clinically diagnosed patients have no genetic diagnosis. Here we report four *MECP2* mutation-negative RTT/RTT-like patients with *de novo* heterozygous novel [NP_001230937.1: p.(Asp248Glu), p.(Cys92Arg) & p.(Pro305Leu)] and previously reported [NP_001230937.1: p.(Thr99Met)] variants in highly-conserved motor domain of kinesin family member 1A (*KIF1A*), which encodes a neuron-specific kinesin-3 motor protein essential for ATP-dependent anterograde axonal transport of synaptic cargo. Along with RTT features, these patients exhibit some clinical features common in children with *KIF1A*-releated disorders. *In*

silico tools predicted these variants to be likely pathogenic, and 3D molecular modelling predicted defective ATP hydrolysis and/or microtubule binding. Using the neurite tip accumulation assay, we demonstrated that all novel variants significantly reduced the ability of motor domains to accumulate in neurite tips of differentiated SH-SY5Y cells. *In-vitro* microtubule gliding assays showed significantly reduced velocities for the variant p.(Asp248Glu) and reduced microtubule binding for the p.(Cys92Arg) and p.(Pro305Leu) variants. Thus, the patient variants result in decreased ability of KIF1A to move along microtubules. This is the first report studying the functional effects of *de novo* missense variants in *KIF1A* in RTT/RTT-like patients.

KEYWORDS

Rett syndrome, *MECP2*, KIF1A, neurite tip accumulation, microtubule, kinesin

GRANT NUMBERS

R01GM070862 and R35GM131744 (KJV), DGE 1256260 (BB)

1 INTRODUCTION

Rett syndrome (RTT; MIM# 312750) is a severe X-linked neurodevelopmental disorder that predominantly affects females and is notable for its progressive nature (Hagberg, Aicardi, Dias, & Ramos, 1983; Rett, 1966). As per the Neul revised diagnostic criteria, RTT patients can be classified broadly into two major categories; classic (or typical RTT) and variant (or atypical RTT) (Neul et al., 2010). Classic RTT patients have normal development until 6 -18 months of age, after which developmental regression occurs, along with the onset of characteristic abnormal hand wringing movements, a hallmark of classic RTT patients. Using the Neul classification, atypical RTT patients do not exhibit all four of the required criteria for classic RTT and must have at least 5 of the 11 supportive criteria. Moreover, there is a further group of RTT-like patients who exhibit some clinical features of RTT, but not enough in order to be classified as either classic or atypical RTT (Ip, Mellios, & Sur, 2018). Up to 97% of classic RTT and 86% of atypical patients have pathogenic variants in Methyl CpG Protein 2 (*MECP2*), and a proportion of individuals with clinical features that overlap with RTT may have variants in cyclin-dependent kinase-like 5 (*CDKL5*), or Forkhead box protein G1 (*FOXG1*) (Mitter et al., 2018; Neul et al., 2014; Schonewolf-Greulich et al., 2019). Due to considerable overlap of

clinical symptoms of RTT patients, especially those who are *MECP2*-mutation negative, with other neurodevelopment disorders, it can be challenging to establish a precise genetic diagnosis. However, recent advances in next generation sequencing (NGS) have identified pathogenic variants in a growing list of genes known to cause intellectual disability, severe epilepsy and/or autistic behaviors where some individuals appear to have a RTT-like clinical picture, thus providing a definitive genetic diagnosis for patients and 'closure' for affected families (Cogliati et al., 2019; Henriksen, Ravn, Paus, von Tetzchner, & Skjeldal, 2018; Iwama et al., 2019; Schonewolf-Greulich et al., 2019; Vidal et al., 2019; Yoo et al., 2017).

There is clear evidence that the RTT-associated gene products MeCP2 and CDKL5 share a common role in regulating microtubule stability and neuronal transport, and that pathogenic mutations in those genes disrupt microtubule dynamics (Barbiero, De Rosa, & Kilstrup-Nielsen, 2019; Delepine, Nectoux, Bahi-Buisson, Chelly, & Bienvenu, 2013; Gold, Lacina, Cantrill, & Christodoulou, 2015)Microtubules are highly dynamic structures that are critical for maintaining correct neuronal architecture and for the transport of cargos via motor proteins such as kinesins (KIFs) and dynein (Kevenaar & Hoogenraad, 2015). KIFs have been implicated in the pathogenesis of RTT suggesting an important molecular role of kinesin family member proteins in this disorder (Gibson et al., 2010; Pelka et al., 2006). They are ATP-dependent molecular motors critical for cytoskeleton organization and in neurons, KIFs are involved in the anterograde transport of synaptic vesicle precursors containing various cargos including brain-derived neurotrophic factor (BDNF; MIM# 113505) (Hirokawa, Noda, Tanaka, & Niwa, 2009; Stanhope & Ross, 2015). BDNF is critical for neuronal differentiation and maintaining synaptic plasticity (Ben Zeev et al., 2009; Li, Calfa, Larimore, & Pozzo-Miller, 2012; Zhou et al., 2006). Interestingly, environmental enrichment in the *Kif1a* $^{+/-}$ mouse induces BDNF-dependent upregulation of kinesin family member 1A (KIF1A; MIM# 601255) and its cargo synaptophysin in the hippocampus, suggesting that BDNF is an upstream regulator of KIF1A (Kondo, Takei, & Hirokawa, 2012). KIF1A, a member of the kinesin-3 family specifically transports cargo containing synaptophysin, synaptotagmin and Ras-associated protein Rab-3A (RAB3A; MIM# 179490). The N-terminal highly conserved motor domain (Supp. Figure S1) has ATP hydrolysis functionality and microtubule binding capacity, which are critical for KIF1A-driven trafficking of specific cargo docked at the C-terminal Pleckstrin-homology (PH) domain. KIF1A dimerizes along the fork-head domain and the neck coil region (Figure 1) (Klopfenstein, Tomishige, Stuurman, & Vale, 2002; Soppina et al., 2014; Tomishige, Klopfenstein, & Vale, 2002).

Pathogenic variants in *KIF1A* were initially associated with autosomal recessive hereditary sensory neuropathy type IIC (HSNIIC; MIM# 614213), autosomal recessive hereditary spastic paraplegia 30 (SPG30; MIM# 610357) and autosomal dominant non-syndromic intellectual disability 9 (MRD9; MIM# 614255) (Erlich et al., 2011; Esmaeeli Nieh et al., 2015; Hotchkiss et al., 2016; Klebe et al., 2012; Krenn et al., 2017; Riviere et al., 2011). More recently, pathogenic variations in *KIF1A* have been recognized to be responsible for a broad spectrum of neurological disorders, broadly termed KIF1A-associated neurological disorders, in which the severity of clinical symptoms largely depends upon the variation introduced and its location in KIF1A (Boyle & Chung, 2017).

HSNIIC-affected individuals have progressive degeneration of sensory neurons, resulting in loss of feeling in extremities, ulceration and ultimately amputation of fingers and toes (Riviere et al., 2011). SPG30 is a slowly progressive spastic paraplegia with unsteady gait, hyperreflexia of the lower limbs (Cheon et al., 2017; Yoshikawa et al., 2019) and the autosomal dominant form also features intellectual disability, speech delay, hypotonia, evolving into hyperreflexia and hypertonia with age. More severely affected individuals have acquired microcephaly, cerebellar atrophy, seizures, scoliosis, contractures, optic nerve atrophy, peripheral neuropathy, ataxia, strabismus, nystagmus, ptosis, facial diplegia and gastroesophageal reflux. In addition, some male patients have been found to have small testes/ penis or cryptorchidism (Boyle & Chung, 2017).

A subset of *KIF1A* patients share some clinical features with RTT including gait abnormalities, hypotonia, scoliosis, seizures and intellectual disability. Recently a *de novo* heterozygous truncating variant [NM_001244008.1: c.275_276insAA; NP_001230937.1: p.(Cys92*)] in the motor domain of KIF1A has been reported in a single female case presenting with classic RTT including developmental delay, microcephaly, lack of independent ambulation and speech, loss of hand skills, hand clapping and mouthing, bruxism, breathing and sleep disturbances (Wang et al., 2019). Interestingly, the well-conserved KIF1A motor domain has been identified as a mutation hotspot (with over 80 reported variants) resulting in a spectrum of phenotypes that overlap with other neurological disorders including RTT. In this study, we report one classic RTT and three RTT-like patients (all *MECP2*-mutation negative) carrying variants in motor domain of KIF1A. We have confirmed the pathogenicity of the variants using a range of *in silico* analyses and *in vitro* functional studies.

2 METHODS

2.1 Subjects and genetic analysis

All methods followed in this study were approved by the human research ethics committee of the participating institutes with written consents obtained from the legal guardians of each participant. The affected participants with variants in the motor domain of KIF1A were recruited from Australia (case 1), Italy (case 2 and 3) and USA (case 4; deceased) and were evaluated as per the revised diagnostic criteria for RTT (Neul et al., 2010). Next generation sequencing (NGS) was performed on genomic DNA isolated from patient blood, and variant filtering was performed as described in the Supplementary Material. The segregation of identified *KIF1A* variants was also confirmed by Sanger sequencing of the proband and the parents.

2.2 *In silico* analysis, variant classification and structural modelling

The pathogenicity of *KIF1A* variants, denoted as per RefSeq NM_001244008.1, was determined using *in silico* prediction tools including PolyPhen-2 (v2.2.2r398; http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), SIFT (v6.2.1; http://sift.jcvi.org/) (Kumar, Henikoff, & Ng, 2009) and MutationTaster (v2.0; www.mutationtaster.org/) (Schwarz, Cooper, Schuelke, & Seelow, 2014). The population frequency was determined using the Genome Aggregation Database (gnomAD) (v2.1.1;http://gnomad.broadinstitute.org/) (Karczewski et al., 2019). All variants were classified according to the guidelines of the American College of Medical Genetics and Genomics (ACMG) (Richards et al., 2015). Amino acid substitutions were assigned a Grantham score (range 0 to 215), with values above 100 generally considered significant (Grantham, 1974). The presence of all these variants was checked in the ClinVar public archive of human genetic variants (https://www.ncbi.nlm.nih.gov/clinvar/; accessed on 16 September 2019) (Landrum et al., 2018).

Since the full length human KIF1A crystal structure is yet to be resolved, the motor domain of *Mus musculus* KIF1A (PDB ID: 2ZFI) was used for *in silico* modelling using HOPE (URL: http://www.cmbi.ru.nl/hope/) (Venselaar, Te Beek, Kuipers, Hekkelman, & Vriend, 2010). The position of residue affected in reported cases is shown on 3D structure images generated using PyMOL (http://www.pymol.org/) molecular visualization software (Schrödinger, 2010). A BLAST search showed 99% identity of the *Mus musculus* KIF1A motor domain

(NP_032466.2) and the motor domain of human KIF1A (NP_001230937.1) (Supplementary Material).

2.3 Constructs and mutagenesis

We used a constitutively-active dimeric KIF1A motor (amino acids 1 – 393) containing the motor (amino acid 1—361), neck linker, and neck coil domains and dimerized via a GCN4 leucine zipper (LZ) (Hammond et al., 2009). KIF1A(1-393)-LZ was tagged at the C-terminus with three tandem mCitrine (mCit) fluorescent proteins [KIF1A(1-393)-3xmCit] for the neurite tip accumulation assay and with mNeonGreen (mNG; Allele Biotech) and AviTag (amino acid sequence GLNDIFEAQKIEWHE) [KIF1A(1-393)-LZ-mNG-AviTag] for the microtubule gliding assay. KIF1A(1-393)-LZ-mNG-AviTag was biotinylated by co-expression with HA-tag labeled bacterial biotin ligase BirA as described (Yue et al., 2018).

Site-directed mutagenesis was used to introduce patient-specific variants into the *KIF1A* constructs (QuikChange lightning mutagenesis kit, Agilent technologies) according to the manufacturer's instructions and using their software for mutagenesis primer design (QuikChange Primer Design, www.agilent.com/genomics/qcpd). Primers are listed in Supp. Table S1.

2.4 Cell culture, transfection, immunostaining and fluorescence microscopy

COS-7 (African green monkey kidney fibroblast, American Type Culture Collection) cells were grown in DMEM media (DMEM with 3.7g/L $NaHCO_3$, 10% (vol/vol) fetal bovine serum (FBS), 1X Pen-Strep) at 37°C in a humidified incubator and 5% CO_2. SH-SY5Y (human neuroblastoma) cells were grown in RPMI media (RPMI with 3.7g/L $NaHCO_3$, 10% (vol/vol) fetal bovine serum (FBS), 1X Pen-Strep) at 37 °C in a humidified incubator and 5% CO_2.

For the neurite tip accumulation assay, 50,000 SH-SY5Y cells were plated on sterile glass coverslips coated with rat collagen (3.3%) and matrigel (2%) (recipe optimised by Dr. Wendy Gold) in RPMI media without serum and antibiotics in a 12 well plate. The following day, cells were transfected with 1µg of KIF1A(1-393)-3xmCit DNA using 5µl of Lipofectamine 2000 (Life Technologies; 11668019) according to the manufacturer's instructions. After 24 hours the media was refreshed to RPMI with 1% serum supplemented with 10µM retinoic acid (RA; Sigma; R2625) to induce neuronal differentiation and the media was refreshed daily for 4 days. The expression of KIF1A motor protein in SH-SY5Y cells did not appear to cause any

qualitative alteration in neurite outgrowth during RA-induced differentiation (data not shown). On day 3, cells were washed with 1X PBS, fixed in 4% (vol/vol) paraformaldehyde (PFA) in PBS for 20 minutes at room temperature before mounting on glass slides using Prolong Gold with DAPI (Life Technologies). Images were acquired using Zeiss AxioVision fluorescence microscope with 40X objective (zoom factor =2). Quantification of the mean florescence intensity (MFI) along the length of the neurites and within the cell bodies of SH-SY5Y cells was measured manually using ImageJ. Mann Whitney test (n=20) was performed on the ratio of average MFI at the neurite length versus the cell body using Prism software (GraphPad). All experiments were performed in triplicate and repeated three times prior to data analysis.

2.5 *In-vitro* microtubule gliding assay

Cell lysates containing biotinylated KIF1A motor proteins were prepared by co-expressing wild type or variant KIF1A(1-393)-LZ-mNG-AviTag constructs together with HA-BirA in COS-7 cells as previously described (Yue et al., 2018). Briefly, COS-7 cells were plated at a density of 200,000 per well in a 6-well plate and transfected with 1 µg of KIF1A(1-393)-LZ-mNG-AviTag and 1 µg HA-BirA DNA using 5 µl of Lipofectamine 2000. Cells were washed, trypsinized and harvested (3000g) at 4^0C 16 hours after transfection. The pellet was washed with PBS and resuspended in ice cold lysis buffer (25 mM HEPES/KOH, pH 7.4, 115 mM potassium acetate, 5 mM sodium acetate, 5 mM $MgCl_2$, 0.5 mM EGTA, and 1% [vol/vol] Triton X-100) supplemented with 1 mM ATP, 1 mM phenylmethylsulphonyl fluoride (PMSF; Sigma; 10837091001), and protease inhibitor cocktail (Thermo Fisher Scientific; A32965) and incubated at 4^0C for 30 minutes. The soluble fraction was collected after centrifugation (20,000g) for 15 mins at 4^0C, snap frozen in single use aliquots, and stored at -80^0C. Biotinylation of KIF1A(1-393)-LZ-mNG-AviTag was confirmed by staining transfected cells with Alexa Fluor® 555 streptavidin (S21381; Thermo Fisher Scientific) and determining coincident staining with green (mNG) fluorescent KIF1A(1-393).

To polymerize microtubules, lyophilized tubulin (Cytoskeleton Inc (USA); T240-A) and rhodamine-labeled tubulin (TL590M-A) were reconstituted to a final concentration of 10 mg/ml using ice-cold general tubulin buffer (80mM PIPES, 2mM $MgCl_2$, 2.5mM EGTA pH 7.0) supplemented with fresh 1mM guanosine triphosphate (GTP; Cytoskeleton; BST06), snap frozen in single use aliquots, and stored at -80^0C. Unlabeled tubulin and labeled tubulin were mixed together in the molar ratio of 10:1, respectively, and incubated at 37^0C for 30 minutes

to induce tubulin polymerization. To maintain the polymerized state, paclitaxel (Cytoskeleton Inc; TXD01) was then added stepwise in increasing concentrations (2µM, 20µM and 200µM) at room temperature with incubation for 10 minutes at each step. The polymerized microtubules were then pelleted through a 40% glycerol cushion at 300,000g for 10 minutes at 25^0C. The pellet was then resuspended using warm general tubulin buffer supplemented with 5% glycerol, 1mM GTP and 20µM paclitaxel.

A 10 µl flow chamber was prepared by attaching two strips of double-sided tape spaced 5 mm apart to a glass slide and mounting a clean #1.5 coverslip (Thermo Fisher Scientific) over the top. Biotinylated KIF1A proteins were immobilized on the attached coverslip by sequential addition of different solutions for 5 minutes each (Yue et al., 2018). Briefly, the flow chamber was first incubated with 1 mg/ml BSA-biotin (Sigma; A8549) followed by blocking solution containing 0.5 mg/ml casein (Sigma; C7078) in BRB80 buffer (80mM PIPES, 1mM $MgCl_2$ and 1mM EGTA) supplemented with 10µM paclitaxel. NeutrAvidin (0.5 mg/ml; Thermo Fischer Scientific; 31000) was then added to the chamber followed by blocking solution before adding 30 µl of COS-7 cell lysates supplemented with 2mM ATP. Then 10µl of motility mix containing taxol-stabilized rhodamine-labeled microtubules, 2mM ATP, 2mM $MgCl_2$ and enzymatic oxygen scavenging system (10mM glucose, 200µg/ml glucose oxidase, 1mM dithiothreitol, 100 µg/ml catalase), to avoid microtubule breaking due to oxidation, was added into the chamber. Finally, the ends of the flow cell were sealed with varnish and microtubule movement was captured by Leica DMIB6000 total internal reflection microscopy (TIRF) microscope fitted with a 100X U PlaApo 1.49 NA oil immersion objective and solid state 488, 561 nm laser line) at room temperature. Signals were collected using EMCCD camera (Andor iXon 897 Ultra). Images were acquired continuously at 50 msec per frame for 4000 frames. Kymographs (width = 3 pixels) were generated by using the multi-kymograph plugin in ImageJ (National Institute of Health) and manually drawing tracks generated from maximum intensity projections as previously published (Yue et al., 2018). The velocity (µm/sec) of microtubule movement was calculated by dividing distance (x-axis) to time (y-axis). Measurements were taken from three independent cell lysate preparations. Plots and statistical analysis using ANOVA and Mann-Whitney test was carried out using Prism software (GraphPad). All data are presented as mean ± standard deviation (SD).

3 RESULTS

In this study we present four unrelated RTT/ RTT-like individuals harboring different *de novo* heterozygous missense variants in *KIF1A*. For these four cases, the clinical features pertaining to RTT are captured in Table 1, and additional clinical details are provided in the supplementary material (Supp. Table S2).

3.1 Variant curation

The four missense variants identified in *KIF1A* are described in Table 2. All the variants were confirmed to be *de novo* heterozygous (Supp. Figure S1) and are located at evolutionarily-conserved sites in the motor domain of KIF1A (Figure 1 and Supp. Figure S2), which is a mutation hot-spot region in *KIF1A*-related disorders (Gabrych, Lau, Niwa, & Silverman, 2019). Several *in silico* tools (MutationTaster, PolyPhen-2 and SIFT) predicted these variants to be disease causing and/or damaging for protein function. Using the Grantham scoring system, the variant c.274T>C; p.(Cys92Arg) was predicted to harbor a major amino acid change with scores of 180, whereas the amino acid changes in c.744C>A; p.(Asp248Glu), c.914C>T; p.(Pro305Leu) and c.296C>T; p.(Thr99Met) were predicted to be minor due to their low Grantham scores of 45, 98 and 81 respectively (Table 2).

All variants were absent in gnomAD (accessed on 16 September 2019) and classified as likely pathogenic using ACMG guidelines. Of the four identified variants, two variants [p.(Asp248Glu) and p.(Cys92Arg)] were absent from ClinVar. For the p.(Pro305Leu) variant, three cases with developmental delay, cerebellar atrophy, ataxia, and eye movement abnormalities were reported in ClinVar. The p.(Thr99Met) is a common recurrent variant and has been reported as a *de novo* heterozygous variant in six unrelated individuals with complex neurological phenotype characterized by moderate to severe development delay and/or intellectual disability, spastic paraparesis, cerebellar atrophy, axonal neuropathy, visual disturbances and/or progressive neurodegeneration (Esmaeeli Nieh et al., 2015; Hamdan et al., 2011; Lee et al., 2015; Okamoto et al., 2014). Moreover, a recent report also found this variant in a female diagnosed with Progressive encephalopathy with Edema, Hypsarrhythmia and Optic atrophy syndrome (PEHO; MIM# 260565) (Langlois et al., 2016).

3.2 Patient variants are located in sequences critical for kinesin enzymatic activity

All four *de novo* variants are located within the kinesin motor domain of KIF1A which harbors the sequences required for ATP-dependent microtubule-based transport. ATP hydrolysis in the motor domain facilitates a step-wise movement of KIF1A along the microtubule network (Atherton et al., 2014; Kikkawa & Hirokawa, 2006; Kikkawa et al., 2001). Three motifs of the kinesin motor domain are involved in coordinating ATP binding and hydrolysis: the phosphate-binding loop (P-loop) with a Walker A fold (GXXXXGK[T/S]), Switch I (loop L9), and Switch II (loop L11) (Kull & Endow, 2013). The variant in case 1 [p.(Asp248Glu)] is present in Switch II and thus, it is predicted to affect nucleotide binding. The wild type (Asp) amino acid coordinates the metal ion (Mg^{2+}) critical for ATP binding, and the increased bulk of the side chain of the introduced amino acid (Glu) may preclude effective nucleotide binding. This is predicted to affect the velocity with which the variant KIF1A would move along the microtubules (Figure 1).

The variant in case 2 [p.(Cys92Arg)] is located in β3 immediately preceding the P-loop. The substitution of the neutral wild-type amino acid (Cys), which is buried in the protein core, with the larger and positively charged variant residue (Arg) is predicted to interfere with KIF1A protein folding and therefore with nucleotide or microtubule binding or both.

Three motifs of the kinesin motor domain have been implicated in microtubule binding: β5-L8, L11-β7, and α4-L12-α5. The variant in case 3 [p.(Pro305Leu)] alters a motif in loop L12 that is evolutionarily-conserved across kinesins for microtubule binding (FIPYRD →FILYRD) (Figure 1). We thus predicted that this variant motor would be ineffective at microtubule binding and motility. The variant in cases 4 [p.(Thr99Met)] affects wild type (Thr) amino acid in the P-loop (nucleotide binding pocket) and hence is predicted to affect the ATP binding to motor domain, thereby altering ATP hydrolysis rate. Structural models published in previous reports suggest that this variant specifically affects interaction between KIF1A motor domain and the γ -phosphate of ATP (Lee et al., 2015).

3.2.1 Patient variants result in defective motility of KIF1A in neuronal SH-SY5Y cells

To examine the effect of patient variants on the efficiency of KIF1A's microtubule-based motility, we utilized the neurite tip accumulation assay. In this assay, kinesin motors that are capable of processive microtubule-based motility accumulate at the tips of neurites in

differentiated neuronal cells (Yang, Bentley, Huang, & Banker, 2016). For this assay, we used a truncated and constitutively active version of KIF1A containing amino acids 1-393 tagged with the fluorescent protein monomeric Citrine (mCit) and we expressed wild-type and mutant versions in differentiated neuronal SH-SY5Y cells.

Consistent with previous work (Hammond et al., 2009; Huang & Banker, 2012), wild-type KIF1A(1-393) accumulated at the end of the growing neurites in differentiated SH-SY5Y cells, indicative of directed movement of KIF1A protein towards the plus ends of the microtubules. However, the KIF1A variant [p.(Asp248Glu), p.Cys92Arg) and p.(Pro305Leu)] proteins did not preferentially accumulate at neurite tips, rather, a significant fraction of the expressed mutant protein was found to be localized in the cell body. The mean fluorescence intensity (MFI) ratio of motor fluorescence in the neurite length compared to the cell body was significantly higher for the wild-type motor protein (2.8±1.4) compared to the patient variants [p.(Asp248Glu) 0.244±0.219; $p<0.0001$], [p.Cys92Arg) 0.118±0.084; $p<0.0001$] and [p.(Pro305Leu) 0.122±0.079; $p<0.0001$] (Figure 2; Supp. Figure S3). Previously published reports have already shown the reduced accumulation of KIF1A motor domain harboring p.(Thr99Met) variant in the distal neurites of rat hippocampal neurons (Hamdan et al., 2011; Lee et al., 2015) These results demonstrate that the variants in KIF1A found in RTT and RTT-like patients have a severe negative impact on the capacity of KIF1A to undergo processive, long-distance transport along the cellular microtubule network.

3.2.2 *In-vitro* microtubule gliding assay reveals impaired microtubule binding

To directly examine the ability of patient variant KIF1A motor proteins to undergo microtubule-based motility, we performed *in-vitro* TIRF-based microtubule gliding assays. In this assay, truncated KIF1A(1-393) motors are immobilized on a slide and their ability to work in concert to "glide" microtubules along the surface is examined by fluorescence microscopy. A unique property of the TIRF microscopy is that it captures the signal from fluorescent microtubules in a limited specimen region immediately adjacent to the adsorbed KIF1A motor protein and the glass coverslip. This exclusive property minimizes any noise coming from the floating microtubules that were not able to bind to the KIF1A motor protein.

For the wild-type KIF1A(1-393) motors, the majority of microtubules (~85-90%) introduced into the flow chamber bound to the immobilized KIF1A and moved along the surface with an average velocity of 0.91 ± 0.09 μm (Figure 3, Supp. Figure S4and Supp. Video 1).

For the patient variant [p.(Asp248Glu)] in switch II, which is predicted to impact the coordination of the divalent cation and ATP, the majority of microtubules introduced into the chamber (~70%) were able to bind to the variant KIF1A motors but moved at a significantly lower velocity (0.005 ± 0.002 μm/s) compared to the wild-type motor protein (Figure 3, Supp. Figure S4 and Supp. Video 2). For the p.(Cys92Arg) and p.(Pro305Leu) variants, very few of the microtubules introduced to the chamber were able to bind to the variant KIF1A motors (~10%) (Figure 3, Supp. Figure S4 and Supp. Video 3, 4). The lack of the microtubule binding was mainly assessed qualitatively by the observation that a significantly smaller proportion of fluorescently labelled microtubules were found to bind to the adsorbed variant as compared to the wild type KIF1A motor protein at the same TIRF interface optimized using wild type protein. Moreover, in the raw videos, the ends of the partially bound microtubules for the p.(Cys92Arg) and p.(Pro305Leu) variants were observed to be flickering before dissociating from the adsorbed protein after a short period of time and moving out of the pre-set interface for TIRF. This phenomenon during the gliding assay suggests that there was impaired microtubule binding, consistent with the hypothesis that these variants alter protein folding and/or microtubule binding. We analyzed the movement of the bound fraction of microtubules and found that these microtubules moved at a much slower velocity [0.014 ± 0.06 μm/s [p.(Cys92Arg); Video 3] and 0.009 ± 0.05 μm/s [p.(Pro305Leu); Video 4] compared to the wild-type motor protein (Figure 3). Nieh et al observed no microtubule motility in case of KIF1A motor protein with p.(Thr99Met) variant which was predicted to affect the ATP binding to the KIF1A motor domain (Esmaeeli Nieh et al., 2015). Overall, these results suggest that patient *KIF1A* variants negatively affect the ability of KIF1A to bind and move along microtubules.

4 DISCUSSION

In this report we describe one classic RTT and three RTT-like patients carrying *de novo* variants in the motor domain of the kinesin-3 motor protein KIF1A. The variants are predicted to be pathogenic by *in silico* and functional analyses. Heterozygous *de novo* variants in *KIF1A*

have previously been associated with HSNIIC, SPG30 and MRD9 as well as a much broader spectrum of *KIF1A*-related disorders (Gabrych et al., 2019).

Here we report a *MECP2*-negative classic RTT female (Case 1) with a novel *de novo* heterozygous p.(Asp248Glu) variant in the motor domain of KIF1A. Characteristic of SPG30, she has progressive lower limb spasticity as well as mild optic nerve atrophy consistent with observations in *KIF1A*-related cases. However, in contrast to *KIF1A*-related disorders, her brain MRI at 12 years of age was normal (Lee et al., 2015; Ohba et al., 2015; Raffa et al., 2017). Recently a classic RTT female patient was reported carrying a *de novo* heterozygous nonsense variant p.(Cys92*) in *KIF1A* with a normal brain MRI (age not specified) (Wang et al., 2019). Our *in vitro* analysis indicates that the p.(Asp248Glu) variant results in a marked decrease in the ability of KIF1A to undergo processive motility to the plus ends of microtubules in the neurite tips of SH-SY5Y cells and a significantly lower velocity of microtubule movement compared to wild-type KIF1A. Together these results suggest that change of Asp248 in switch II of KIF1A results in compromised ATP hydrolysis/cycling rates.

The RTT-like cases [case 2 with p.(Cys92Arg) and case 3 with p.(Pro305Leu)] also had features overlapping with the other *KIF1A*-related phenotypes. Their RTT-like features included microcephaly, global developmental delay, seizures, stereotypic hand movements, gait abnormalities, abnormal muscle tone and scoliosis. Our functional assays provide strong evidence that these RTT-like variants negatively impact KIF1A function. The variant KIF1A proteins, p.(Cys92Arg) and p.(Pro305Leu) were unable to accumulate at the neurite tips in differentiated SH-SY5Y cells, indicating a lack of processive motility along microtubules for these KIF1A variant proteins. While the p.(Cys92Arg) variant was able to bind to microtubules in the gliding assay, the velocity of movement was dramatically decreased. The p.(Pro305Leu) variant in L12 was predicated to impact the ability of KIF1A to bind to microtubules and indeed, the variant KIF1A protein showed a significant reduction in its ability to bind to microtubules in the *in vitro* microtubule gliding assay.

The RTT-like case 4 with p.(Thr99Met) variant in *KIF1A* shared common clinical features with the previously reported patients with this variant such as abnormal muscle tone, seizures, intellectual disability, gait abnormalities, optic nerve abnormalities (Hamdan et al., 2011; Langlois et al., 2016; Lee et al., 2015). However, in contrast to already published patients, our case 4 exhibited characteristic RTT features such as period of regression followed by

stabilization as well as characteristic stereotypic hand movements. Already published functional data showed a significant decrease in the accumulation of variant KIF1A protein at the distal ends of primary rat hippocampal neurons as well as no microtubule motility using the *in-vitro* microtubule gliding assay. This is consistent with molecular modelling results predicting loss in interaction between the KIF1A motor domain and ATP, which is crucial for the movement of kinesin head on the microtubule structure (Esmaeeli Nieh et al., 2015; Hamdan et al., 2011; Lee et al., 2015). Together these findings imply that the KIF1A variant proteins are defective for their ability to bind to and move along microtubules, and support the conclusion that they are pathogenic.

In support of these molecular findings, it is known that KIF1A is regulated by BDNF which is a well-known target of MeCP2 (Li & Pozzo-Miller, 2014). Several reports have uncovered BDNF defects in RTT mouse models and patients at the genetic as well as functional levels (Ben Zeev et al., 2009; Li et al., 2012; Sampathkumar et al., 2016; Zhou et al., 2006). BDNF has also been shown to be an upstream regulator of KIF1A and its specific cargo such as synaptophysin (Kondo et al., 2012). Importantly, variants in *KIF1A* have been previously described to cause intellectual disability, highlighting the potential pathogenic relevance of this gene in RTT. Thus, we suggest that defects in KIF1A could contribute towards the pathophysiology of RTT, but the exact molecular mechanism still needs to be explored.

In conclusion, we have identified four *de novo* heterozygous variants in *KIF1A* in four RTT/RTT-like patients who share common clinical features with the other *KIF1A*-related disorders. We recommend that genetic testing for *KIF1A* should be considered in *MECP2* negative RTT/RTT-like patients, and that the clinical spectrum of *KIF1A*-related disorders should be expanded to include RTT and RTT-like clinical phenotypes.

5 CONFLICT OF INTEREST

The authors declare no conflict of interest.

6 DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

7 ACKNOWLEDGMENTS

We thank all the family members for taking part in this study. The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. SK was the recipient of Research Training Program scholarship. Work in the laboratory of KJV is supported by grants (R01GM070862 and R35GM131744) from the United States' National Institutes of Health. BB was supported by a Graduate Research Fellowship from the National Science Foundation under grant no. DGE 1256260, the Rackham Pre-doctoral Fellowship, and the Endowment for the Development of Graduate Education Student Fellowship

8 REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

Atherton, J., Farabella, I., Yu, I. M., Rosenfeld, S. S., Houdusse, A., Topf, M., & Moores, C. A. (2014). Conserved mechanisms of microtubule-stimulated ADP release, ATP binding, and force generation in transport kinesins. *Elife, 3*, e03680. doi:10.7554/eLife.03680

Barbiero, I., De Rosa, R., & Kilstrup-Nielsen, C. (2019). Microtubules: A Key to Understand and Correct Neuronal Defects in CDKL5 Deficiency Disorder? *Int J Mol Sci, 20*(17). doi:10.3390/ijms20174075

Ben Zeev, B., Bebbington, A., Ho, G., Leonard, H., de Klerk, N., Gak, E., . . . Christodoulou, J. (2009). The common BDNF polymorphism may be a modifier of disease severity in Rett syndrome. *Neurology, 72*(14), 1242-1247. doi:10.1212/01.wnl.0000345664.72220.6a

Boyle, L., & Chung, W. (2017). *Expanding the natural history of KIF1A associated neurological disorders (KAND)*.

Cheon, C. K., Lim, S. H., Kim, Y. M., Kim, D., Lee, N. Y., Yoon, T. S., . . . Lee, J. R. (2017). Autosomal dominant transmission of complicated hereditary spastic paraplegia due to a dominant negative mutation of KIF1A, SPG30 gene. *Sci Rep, 7*(1), 12527. doi:10.1038/s41598-017-12999-9

Cogliati, F., Giorgini, V., Masciadri, M., Bonati, M. T., Marchi, M., Cracco, I., . . . Russo, S. (2019). Pathogenic Variants in STXBP1 and in Genes for GABAa Receptor Subunities Cause Atypical Rett/Rett-like Phenotypes. *Int J Mol Sci, 20*(15). doi:10.3390/ijms20153621

Delepine, C., Nectoux, J., Bahi-Buisson, N., Chelly, J., & Bienvenu, T. (2013). MeCP2 deficiency is associated with impaired microtubule stability. *FEBS Lett, 587*(2), 245-253. doi:10.1016/j.febslet.2012.11.033

Erlich, Y., Edvardson, S., Hodges, E., Zenvirt, S., Thekkat, P., Shaag, A., . . . Elpeleg, O. (2011). Exome sequencing and disease-network analysis of a single family implicate a mutation in KIF1A in hereditary spastic paraparesis. *Genome Res, 21*(5), 658-664. doi:10.1101/gr.117143.110

Esmaeeli Nieh, S., Madou, M. R., Sirajuddin, M., Fregeau, B., McKnight, D., Lexa, K., . . . Sherr, E. H. (2015). De novo mutations in KIF1A cause progressive encephalopathy and brain atrophy. *Ann Clin Transl Neurol, 2*(6), 623-635. doi:10.1002/acn3.198

Gabrych, D. R., Lau, V. Z., Niwa, S., & Silverman, M. A. (2019). Going Too Far Is the Same as Falling Short(dagger): Kinesin-3 Family Members in Hereditary Spastic Paraplegia. *Front Cell Neurosci, 13*, 419. doi:10.3389/fncel.2019.00419

Gibson, J. H., Slobedman, B., K, N. H., Williamson, S. L., Minchenko, D., El-Osta, A., . . . Christodoulou, J. (2010). Downstream targets of methyl CpG binding protein 2 and their abnormal expression in the frontal cortex of the human Rett syndrome brain. *BMC Neurosci, 11*, 53. doi:10.1186/1471-2202-11-53

Gold, W. A., Lacina, T. A., Cantrill, L. C., & Christodoulou, J. (2015). MeCP2 deficiency is associated with reduced levels of tubulin acetylation and can be restored using HDAC6 inhibitors. *J Mol Med (Berl), 93*(1), 63-72. doi:10.1007/s00109-014-1202-x

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. *Science, 185*(4154), 862-864. doi:10.1126/science.185.4154.862

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

Hamdan, F. F., Gauthier, J., Araki, Y., Lin, D. T., Yoshizawa, Y., Higashi, K., . . . Michaud, J. L. (2011). Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet, 88*(3), 306-316. doi:10.1016/j.ajhg.2011.02.001

Hammond, J. W., Cai, D., Blasius, T. L., Li, Z., Jiang, Y., Jih, G. T., . . . Verhey, K. J. (2009). Mammalian Kinesin-3 motors are dimeric in vivo and move by processive motility upon release of autoinhibition. *PLoS Biol, 7*(3), e72. doi:10.1371/journal.pbio.1000072

Henriksen, M. W., Ravn, K., Paus, B., von Tetzchner, S., & Skjeldal, O. H. (2018). De novo mutations in SCN1A are associated with classic Rett syndrome: a case report. *BMC Med Genet, 19*(1), 184. doi:10.1186/s12881-018-0700-z

Hirokawa, N., Noda, Y., Tanaka, Y., & Niwa, S. (2009). Kinesin superfamily motor proteins and intracellular transport. *Nat Rev Mol Cell Biol, 10*(10), 682-696. doi:10.1038/nrm2774

Hotchkiss, L., Donkervoort, S., Leach, M. E., Mohassel, P., Bharucha-Goebel, D. X., Bradley, N., . . . Bonnemann, C. G. (2016). Novel De Novo Mutations in KIF1A as a Cause of Hereditary Spastic Paraplegia With Progressive Central Nervous System Involvement. *J Child Neurol, 31*(9), 1114-1119. doi:10.1177/0883073816639718

Huang, C. F., & Banker, G. (2012). The translocation selectivity of the kinesins that mediate neuronal organelle transport. *Traffic, 13*(4), 549-564. doi:10.1111/j.1600-0854.2011.01325.x

Ip, J. P. K., Mellios, N., & Sur, M. (2018). Rett syndrome: insights into genetic, molecular and circuit mechanisms. *Nat Rev Neurosci, 19*(6), 368-382. doi:10.1038/s41583-018-0006-3

Iwama, K., Mizuguchi, T., Takeshita, E., Nakagawa, E., Okazaki, T., Nomura, Y., . . . Matsumoto, N. (2019). Genetic landscape of Rett syndrome-like phenotypes revealed by whole exome sequencing. *J Med Genet, 56*(6), 396-407. doi:10.1136/jmedgenet-2018-105775

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., . . . MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. 531210. doi:10.1101/531210 %J bioRxiv

Kevenaar, J. T., & Hoogenraad, C. C. (2015). The axonal cytoskeleton: from organization to function. *Front Mol Neurosci, 8*, 44. doi:10.3389/fnmol.2015.00044

Kikkawa, M., & Hirokawa, N. (2006). High-resolution cryo-EM maps show the nucleotide binding pocket of KIF1A in open and closed conformations. *EMBO J, 25*(18), 4187-4194. doi:10.1038/sj.emboj.7601299

Kikkawa, M., Sablin, E. P., Okada, Y., Yajima, H., Fletterick, R. J., & Hirokawa, N. (2001). Switch-based mechanism of kinesin motors. *Nature, 411*(6836), 439-445. doi:10.1038/35078000

Klebe, S., Lossos, A., Azzedine, H., Mundwiller, E., Sheffer, R., Gaussen, M., . . . Stevanin, G. (2012). KIF1A missense mutations in SPG30, an autosomal recessive spastic paraplegia: distinct phenotypes according to the nature of the mutations. *Eur J Hum Genet, 20*(6), 645-649. doi:10.1038/ejhg.2011.261

Klopfenstein, D. R., Tomishige, M., Stuurman, N., & Vale, R. D. (2002). Role of phosphatidylinositol(4,5)bisphosphate organization in membrane transport by the Unc104 kinesin motor. *Cell, 109*(3), 347-358. doi:10.1016/s0092-8674(02)00708-0

Kondo, M., Takei, Y., & Hirokawa, N. (2012). Motor protein KIF1A is essential for hippocampal synaptogenesis and learning enhancement in an enriched environment. *Neuron, 73*(4), 743-757. doi:10.1016/j.neuron.2011.12.020

Krenn, M., Zulehner, G., Hotzy, C., Rath, J., Stogmann, E., Wagner, M., . . . Zimprich, F. (2017). Hereditary spastic paraplegia caused by compound heterozygous mutations outside the motor domain of the KIF1A gene. *Eur J Neurol, 24*(5), 741-747. doi:10.1111/ene.13279

Kull, F. J., & Endow, S. A. (2013). Force generation by kinesin and myosin cytoskeletal motor proteins. *J Cell Sci, 126*(Pt 1), 9-19. doi:10.1242/jcs.103911

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc, 4*(7), 1073-1081. doi:10.1038/nprot.2009.86

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., . . . Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res, 46*(D1), D1062-D1067. doi:10.1093/nar/gkx1153

Langlois, S., Tarailo-Graovac, M., Sayson, B., Drogemoller, B., Swenerton, A., Ross, C. J., . . . van Karnebeek, C. D. (2016). De novo dominant variants affecting the motor domain of KIF1A are a cause of PEHO syndrome. *Eur J Hum Genet, 24*(6), 949-953. doi:10.1038/ejhg.2015.217

Lee, J. R., Srour, M., Kim, D., Hamdan, F. F., Lim, S. H., Brunel-Guitton, C., . . . Michaud, J. L. (2015). De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. *Hum Mutat, 36*(1), 69-78. doi:10.1002/humu.22709

Li, W., Calfa, G., Larimore, J., & Pozzo-Miller, L. (2012). Activity-dependent BDNF release and TRPC signaling is impaired in hippocampal neurons of Mecp2 mutant mice. *Proc Natl Acad Sci U S A, 109*(42), 17087-17092. doi:10.1073/pnas.1205271109

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

Mitter, D., Pringsheim, M., Kaulisch, M., Plumacher, K. S., Schroder, S., Warthemann, R., . . . Brockmann, K. (2018). FOXG1 syndrome: genotype-phenotype association in 83 patients with FOXG1 variants. *Genet Med, 20*(1), 98-108. doi:10.1038/gim.2017.75

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

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

Ohba, C., Haginoya, K., Osaka, H., Kubota, K., Ishiyama, A., Hiraide, T., . . . Matsumoto, N. (2015). De novo KIF1A mutations cause intellectual deficit, cerebellar atrophy, lower limb spasticity and visual disturbance. *J Hum Genet, 60*(12), 739-742. doi:10.1038/jhg.2015.108

Okamoto, N., Miya, F., Tsunoda, T., Yanagihara, K., Kato, M., Saitoh, S., . . . Kosaki, K. (2014). KIF1A mutation in a patient with progressive neurodegeneration. *J Hum Genet, 59*(11), 639-641. doi:10.1038/jhg.2014.80

Pelka, G. J., Watson, C. M., Radziewic, T., Hayward, M., Lahooti, H., Christodoulou, J., & Tam, P. P. (2006). Mecp2 deficiency is associated with learning and cognitive deficits and altered gene activity in the hippocampal region of mice. *Brain, 129*(Pt 4), 887-898. doi:10.1093/brain/awl022

Raffa, L., Matton, M. P., Michaud, J., Rossignol, E., Decarie, J. C., & Ospina, L. H. (2017). Optic nerve hypoplasia in a patient with a de novo KIF1A heterozygous mutation. *Can J Ophthalmol, 52*(5), e169-e171. doi:10.1016/j.jcjo.2017.02.021

Rett, A. (1966). [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien Med Wochenschr, 116*(37), 723-726.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Riviere, J. B., Ramalingam, S., Lavastre, V., Shekarabi, M., Holbert, S., Lafontaine, J., . . . Rouleau, G. A. (2011). KIF1A, an axonal transporter of synaptic vesicles, is mutated in hereditary sensory and autonomic neuropathy type 2. *Am J Hum Genet, 89*(2), 219-230. doi:10.1016/j.ajhg.2011.06.013

Sampathkumar, C., Wu, Y. J., Vadhvani, M., Trimbuch, T., Eickholt, B., & Rosenmund, C. (2016). Loss of MeCP2 disrupts cell autonomous and autocrine BDNF signaling in mouse glutamatergic neurons. *Elife, 5*. doi:10.7554/eLife.19374

Schonewolf-Greulich, B., Bisgaard, A. M., Moller, R. S., Duno, M., Brondum-Nielsen, K., Kaur, S., . . . Tumer, Z. (2019). Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin Genet, 95*(2), 221-230. doi:10.1111/cge.13153

Schrödinger. (2010). The PyMOL Molecular Graphics System.

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods, 11*(4), 361-362. doi:10.1038/nmeth.2890

Soppina, V., Norris, S. R., Dizaji, A. S., Kortus, M., Veatch, S., Peckham, M., & Verhey, K. J. (2014). Dimerization of mammalian kinesin-3 motors results in superprocessive

motion. *Proc Natl Acad Sci U S A, 111*(15), 5562-5567. doi:10.1073/pnas.1400759111

Stanhope, K., & Ross, J. (2015). Microtubules, MAPs, and Motor Proteins. In J. Ross & W. Marshall (Eds.), *Methods in Cell Biology* (Vol. 128, pp. 23-37).

Tomishige, M., Klopfenstein, D. R., & Vale, R. D. (2002). Conversion of Unc104/KIF1A kinesin into a processive motor after dimerization. *Science, 297*(5590), 2263-2267. doi:10.1126/science.1073386

Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics, 11*, 548. doi:10.1186/1471-2105-11-548

Vidal, S., Brandi, N., Pacheco, P., Maynou, J., Fernandez, G., Xiol, C., . . . Armstrong, J. (2019). The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome. *Eur J Paediatr Neurol, 23*(4), 609-620. doi:10.1016/j.ejpn.2019.04.006

Wang, J., Zhang, Q., Chen, Y., Yu, S., Wu, X., & Bao, X. (2019). Rett and Rett-like syndrome: Expanding the genetic spectrum to KIF1A and GRIN1 gene. *Mol Genet Genomic Med*, e968. doi:10.1002/mgg3.968

Yang, R., Bentley, M., Huang, C. F., & Banker, G. (2016). Analyzing kinesin motor domain translocation in cultured hippocampal neurons. *Methods Cell Biol, 131*, 217-232. doi:10.1016/bs.mcb.2015.06.021

Yoo, Y., Jung, J., Lee, Y. N., Lee, Y., Cho, H., Na, E., . . . Choi, M. (2017). GABBR2 mutations determine phenotype in rett syndrome and epileptic encephalopathy. *Ann Neurol, 82*(3), 466-478. doi:10.1002/ana.25032

Yoshikawa, K., Kuwahara, M., Saigoh, K., Ishiura, H., Yamagishi, Y., Hamano, Y., . . . Kusunoki, S. (2019). The novel de novo mutation of KIF1A gene as the cause for Spastic paraplegia 30 in a Japanese case. *eNeurologicalSci, 14*, 34-37. doi:10.1016/j.ensci.2018.11.026

Yue, Y., Blasius, T. L., Zhang, S., Jariwala, S., Walker, B., Grant, B. J., . . . Verhey, K. J. (2018). Altered chemomechanical coupling causes impaired motility of the kinesin-4 motors KIF27 and KIF7. *J Cell Biol, 217*(4), 1319-1334. doi:10.1083/jcb.201708179

Zhou, Z., Hong, E. J., Cohen, S., Zhao, W. N., Ho, H. Y., Schmidt, L., . . . Greenberg, M. E. (2006). Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron, 52*(2), 255-269. doi:10.1016/j.neuron.2006.09.037

9 FIGURE LEGENDS

Figure 1: Schematic representing KIF1A domain structure, evolutionary conservation at variant site and 3D molecular modelling.

A) Domain structure of full length KIF1A containing N-terminal motor domain (MD: 1-361 aa), neck coil (NC: 365-397 aa), coiled coil domains [(CC1: 429-462 aa), (CC2; 622-681 aa) and (CC3: 801-822 aa)], forkhead-associated domain (FH: 516-572 aa), and pleckstrin homology domain (PH: 1676-1774 aa)]. B) Representation showing location of four patient variants in four RTT/ RTT-like cases in motor domain of KIF1A. The variants NP_001230937.1:p.(Asp248Glu), NP_001230937.1:p.(Cys92Arg) and NP_001230937.1:p.(Pro305Leu) are novel (in black), whereas NP_001230937.1:p.(Thr99Met) is previously reported (in red). Multiple sequence alignment underneath each variant shows the sequence conservation at each variant site. Sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_001230937.1), *M musculus* (NP_032466.2), *R norvegicus* (XP_017452404.1), *C elegans* (NP_001022041.2) and *D melanogaster* (NP_001246373.1). C) Schematic structures representing location of variants on crystal structure of KIF1A motor domain before Mg ion release (PDB: 2ZFI) using PyMOL software. The P-loop, switch I and switch II, P-loop are indicated in light magenta, aquamarine and yellow-orange respectively and rest of motor domain is represented in grey. The ligand Adenosine-5'-diphosphate (ADP) is shown by a stick model in blue and magnesium ion is represented as orange sphere. The residue affected in reported cases are represented as red circles and red labels.

Figure 2: Neurite tip accumulation in differentiated SH-SY5Y cells.

A) Representative images showing the localization of rat KIF1A(1-393)-3xmCit motors (green) in differentiated SH-SY5Y cells (Nucleus/ DAPI = blue). Whereas the wild-type motor accumulated in neurite tips (white arrows), the variant proteins remained near the cell body. B) Quantification of the mean florescence intensity (MFI) at the neurite length versus the cell body for the wild-type and three pathogenic *KIF1A* patient variants. Data are from three independent experiments (n=20 cells for each experiment) and are plotted as mean ± standard deviation. **** represents $p<0.0001$, Mann-Whitney test. Scale bar (white line, bottom right) = 10µm.

Figure 3: *in vitro* microtubule gliding driven by wild-type and variant KIF1A motors.
A) Schematic of the microtubule gliding assay in a flow chamber modified from (Yue et al., 2018). The biotinylated KIF1A(1-393)-LZ-mNG-AviTag motor proteins expressed in COS-7 cells were immobilized on a NeutrAvidin-coated glass coverslip. The motility mix containing rhodamine-labelled microtubules, ATP, and oxygen-scavenging system was then added into the flow chamber. Time lapse images were captured using the TIRF microscope. B) Representative kymographs (distance versus time graphs) for rhodamine-labeled microtubule movement over "lawns" of wild-type or variant KIF1A proteins. Time (seconds) is on y-axis and distance travelled (μm) is on x-axis. The angled line represents movement and straight line signifies stationary microtubules. C) Quantification of the percentage of microtubules that were stationary versus motile. D) Violin plot representing average velocity (μm/sec) of microtubules moved by KIF1A(1-393) proteins. Black = violin plot, blue = each data point, red diamond with center = mean, red line = +/- standard deviation. The average velocity for every variant KIF1A motor domain protein was significantly reduced compared to the wild type protein. Data are analyzed from three independent experiments (n=50 microtubules each) and is plotted as mean ± standard deviation. **** represents $p<0.0001$, Mann-Whitney test.

TABLES

Table 1: Assessment of all cases as per the revised diagnostic criteria (Neul et al 2010)

Category	Case 1	Case 2	Case 3	Case 4 (deceased)
Age (as of 2019)	19 years	10 years	6 years	5 years 9 month
Gender	Female	Female	Male	Female
Ancestry	Vietnamese	Italian	Italian	French Canadian
Main criteria				
Partial/ complete loss of acquired purposeful hand skills	Yes	No	No	No
Partial/ complete loss of acquired spoken language	Yes	No	No	No
Gait abnormalities: Impaired/absence of ability	Yes	Yes	Yes	Yes
Stereotypic hand movements†	Yes	Yes	Yes	No
Supportive criteria for atypical RTT				
Breathing disturbances when awake	Yes	No	No	No
Bruxism when awake	Yes	No	No	Yes
Impaired sleep pattern	No	No	No	No
Abnormal muscle tone	Yes	Yes	Yes	Yes
Peripheral vasomotor disturbances	No	No	No	No
Scoliosis/kyphosis	No	Yes	No	Yes
Growth retardation	Yes	No	Yes	No
Small cold hands and feet	No	No	No	Yes
Inappropriate laughing/screaming spells	No	No	Yes	No
Diminished response to pain	No	No	No	Yes
Intense eye communication - ‘‘eye pointing’’	Yes	No	No	No
Required for classic (typical) RTT				
A period of regression followed by recovery or stabilization	Yes	No	No	Yes
All main criteria and all exclusion criteria	Yes	No	No	No
Supportive criteria (not required, often seen in classic RTT)	Yes	No	No	Yes
Exclusion criteria				
Brain injury‡	No	No	No	No
Grossly abnormal psychomotor development (first 6 months)	No	Yes	No	Yes
Required for atypical or variant RTT				
A period of regression followed by recovery or stabilization	Yes	No	No	Yes
At least 2 of the 4 main criteria	Yes (4/4)	Yes (2/4)	Yes (2/4)	No (1/4)
5 out of 11 supportive criteria	Yes (5/11)	No (2/11)	No (3/11)	Yes (5/11)
Classification	**Classic RTT**	**RTT-like**	**RTT-like**	**RTT-like**

†: hand wringing/squeezing, clapping/ tapping, mouthing and washing/rubbing automatisms; ‡: secondary to trauma (peri- or postnatally), neurometabolic disease, or severe infection that causes neurological problems; NA= Not available

Table 2: Variant analysis for RTT/RTT-like cases

Category		Case 1	Case 2	Case 3	Case 4 (deceased)
Genetic screening method		trio; WGS	trio; WES	trio; targeted sequencing	Singleton; targeted sequencing
Variant		chr2:241723210G>T; c.744C>A; p.(Asp248Glu)	chr2:241727557A>G; c.274T>C; p.(Cys92Arg)	chr2:241715312G>A; c.914C>T; p.(Pro305Leu)	chr2:241727535G>A; c.296C>T; (Thr99Met)
Zygosity		Heterozygous	Heterozygous	Heterozygous	Heterozygous
Segregation		*de novo*	*de novo*	*de novo*	*de novo*
Conserved site		Yes	Yes	Yes	Yes
Novel or reported		Novel	Novel	Reported	Reported
Affected exon		Exon 8	Exon 4	Exon 11	Exon 4
Amino acid change (Grantham distance)		Minor (score: 45)	Major (score: 180)	Minor (score: 98)	Minor (score: 81)
population database (gnomAD)		No	No	No	No
Location of variant		Motor domain	Motor domain	Motor domain	Motor domain
***in silico* tools**	**Mutation Taster#**	Disease causing (p=0.999)	Disease causing (0.999)	Disease causing (0.999)	Disease causing (0.999)
	PolyPhen-2##	Probably damaging (p=0.998)	Probably damaging (p=1.0)	Probably damaging (p=1.0)	Probably damaging (p=1.0)
	SIFT###	Damaging (score = 0.001)	Damaging (score = 0.000)	Damaging (score = 0.001)	Damaging (score = 0.000)
ClinVar		No	No	Yes	Yes
ACMG classification		Likely pathogenic	Likely pathogenic	Likely pathogenic	Likely pathogenic

RefSeq ID: NM_001244008.1 (longest isoform): Position: GRCh37 chr2:241,653,181-241,759,725, size: 106,545 and total exon count: 49; NCBI Protein ID: NP_001230937.1 (longest isoform); Ensembl transcript ID: ENST00000320389. WGS = whole genome sequencing, WES = whole exome sequencing, ACMG= American College of Medical Genetics, gnomAD = Genome Aggregation Database.

#: v2.0. URL = www.mutationtaster.org/

##: v2.2.2r398. URL = http://genetics.bwh.harvard.edu/pph2/. Score range: 0.0 (tolerated) to 1.0 (deleterious)

###: v6.2.1. URL = http://sift.jcvi.org/. Score range: 0.0 (deleterious) to 1.0 (benign). Prediction cut-off = 0.05.

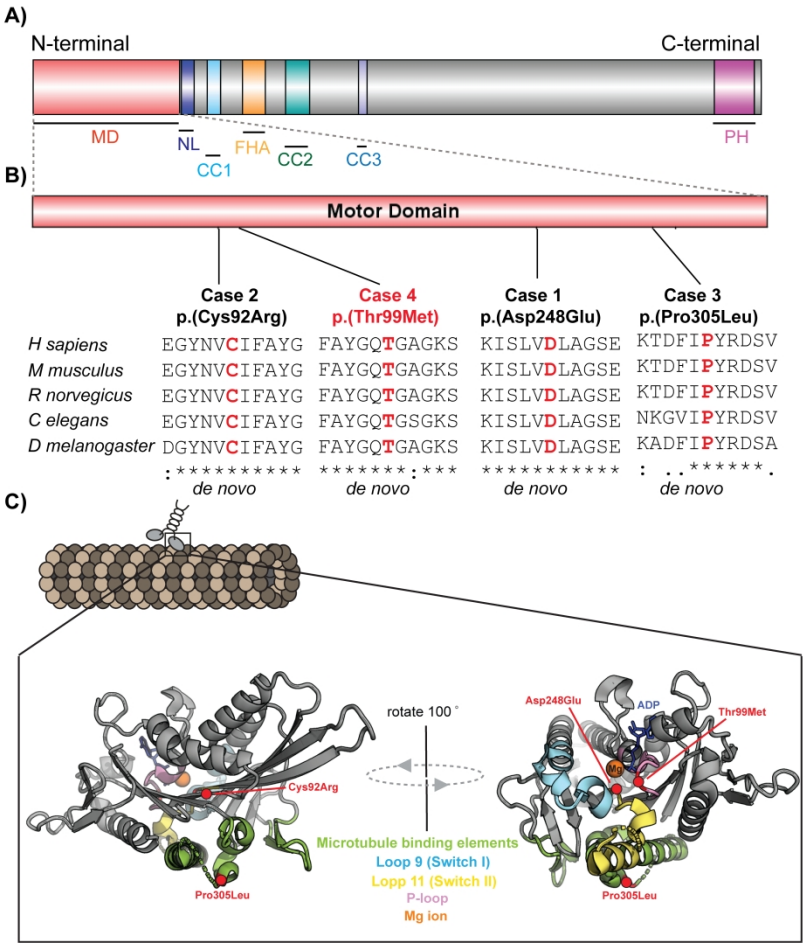


Figure 1: Schematic representing KIF1A domain structure, evolutionary conservation at variant site and 3D molecular modelling.
A) Domain structure of full length KIF1A containing N-terminal motor domain (MD: 1-361 aa), neck coil (NC: 365-397 aa), coiled coil domains [(CC1: 429-462 aa), (CC2; 622-681 aa) and (CC3: 801-822 aa)], forkhead-associated domain (FH: 516-572 aa), and pleckstrin homology domain (PH: 1676-1774 aa)]. B) Representation showing location of four patient variants in four RTT/ RTT-like cases in motor domain of KIF1A. The variants NP_001230937.1:p.(Asp248Glu), NP_001230937.1:p.(Cys92Arg) and NP_001230937.1:p.(Pro305Leu) are novel (in black), whereas NP_001230937.1:p.(Thr99Met) is previously reported (in red). Multiple sequence alignment underneath each variant shows the sequence conservation at each variant site. Sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of H sapiens (NP_001230937.1), M musculus (NP_032466.2), R norvegicus (XP_017452404.1), C elegans (NP_001022041.2) and D melanogaster (NP_001246373.1). C) Schematic structures representing location of variants on crystal structure of KIF1A motor domain before Mg ion release (PDB: 2ZFI) using PyMOL software. The P-loop, switch I and switch II, P-loop are indicated in light magenta, aquamarine and yellow-orange respectively and rest of motor domain is represented in grey. The ligand Adenosine-5'-diphosphate (ADP) is shown by a stick model in blue and magnesium ion is represented as orange sphere. The residue affected in reported cases are represented as red circles and red labels.

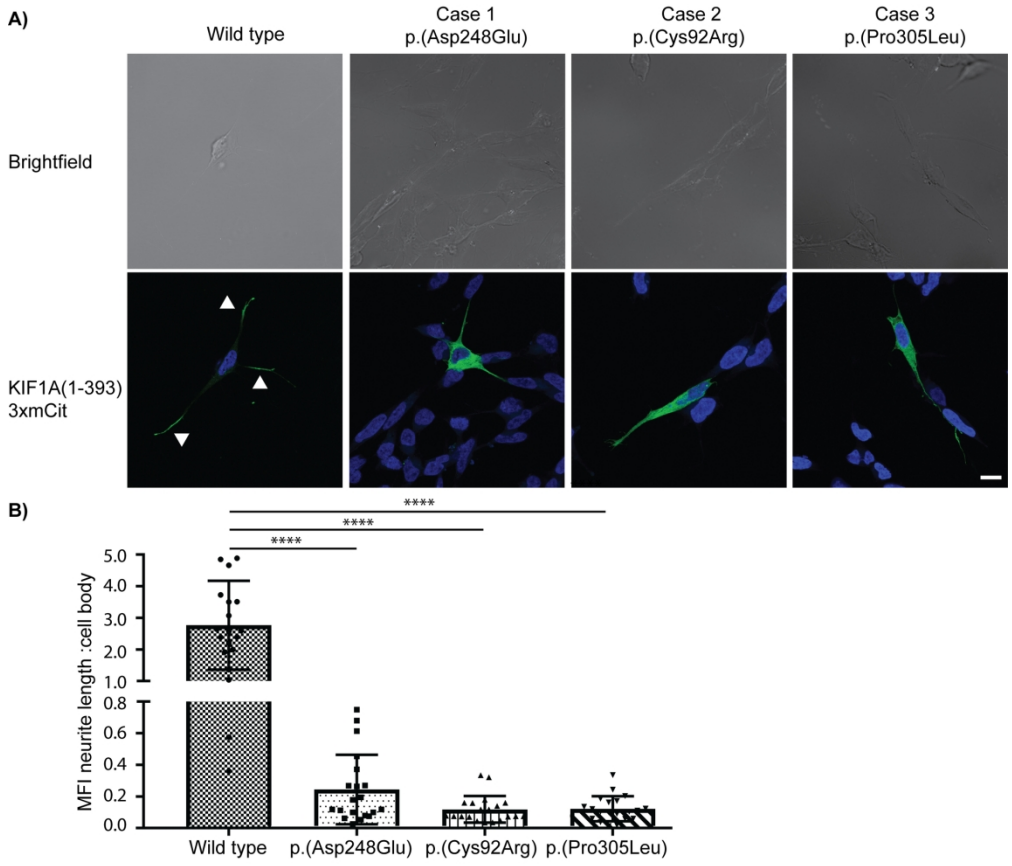


Figure 2: Neurite tip accumulation in differentiated SH-SY5Y cells.
A) Representative images showing the localization of rat KIF1A(1-393)-3xmCit motors (green) in differentiated SH-SY5Y cells (Nucleus/ DAPI = blue). Whereas the wild-type motor accumulated in neurite tips (white arrows), the variant proteins remained near the cell body. B) Quantification of the mean florescence intensity (MFI) at the neurite length versus the cell body for the wild-type and three pathogenic KIF1A patient variants. Data are from three independent experiments (n=20 cells for each experiment) and are plotted as mean ± standard deviation. **** represents $p<0.0001$, Mann-Whitney test. Scale bar (white line, bottom right) = 10µm.

185x163mm (300 x 300 DPI)

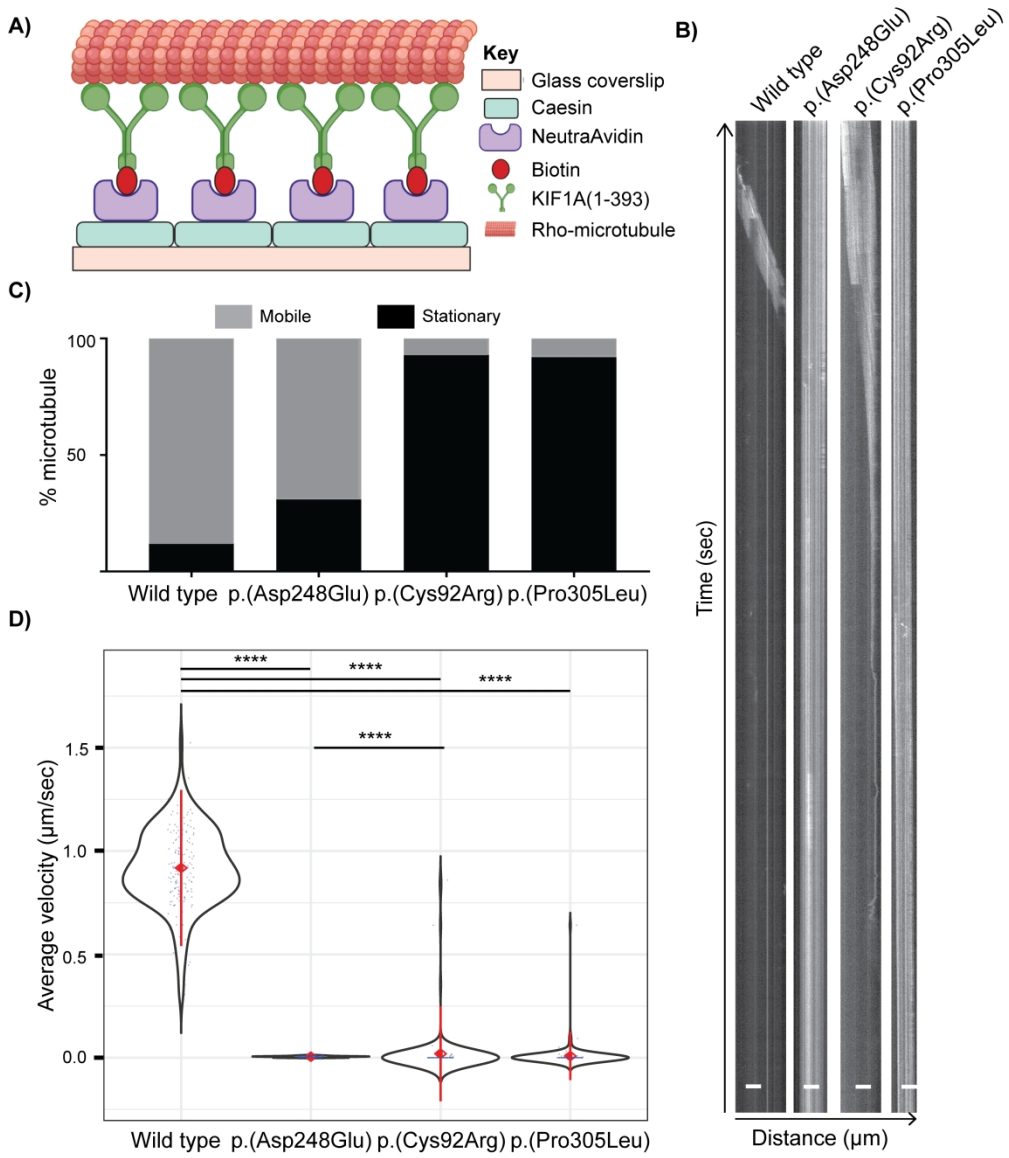


Figure 3: in vitro microtubule gliding driven by wild-type and variant KIF1A motors.
A) Schematic of the microtubule gliding assay in a flow chamber modified from (Yue et al., 2018). The biotinylated KIF1A(1-393)-LZ-mNG-AviTag motor proteins expressed in COS-7 cells were immobilized on a NeutrAvidin-coated glass coverslip. The motility mix containing rhodamine-labelled microtubules, ATP, and oxygen-scavenging system was then added into the flow chamber. Time lapse images were captured using the TIRF microscope. B) Representative kymographs (distance versus time graphs) for rhodamine-labeled microtubule movement over "lawns" of wild-type or variant KIF1A proteins. Time (seconds) is on y-axis and distance travelled (µm) is on x-axis. The angled line represents movement and straight line signifies stationary microtubules. C) Quantification of the percentage of microtubules that were stationary versus motile. D) Violin plot representing average velocity (µm/sec) of microtubules moved by KIF1A(1-393) proteins. Black = violin plot, blue = each data point, red diamond with center = mean, red line = +/- standard deviation. The average velocity for every variant KIF1A motor domain protein was significantly reduced compared to the wild type protein. Data are analyzed from three independent experiments (n=50 microtubules each) and is plotted as mean ± standard deviation. **** represents $p<0.0001$, Mann-Whitney

Supplementary Information

Phenotypic spectrum of *de novo* missense variants in kinesin family member 1A (*KIF1A*): expansion to include Rett syndrome and Rett-like patients

Simranpreet Kaur[1,2*], Nicole J. Van Bergen[1,2*], Kristen J Verhey[3], Cameron J. Nowell[4], Breane Budaitis[5], Yang Yue[3], Carolyn Ellaway[6,7], Nicola Brunetti-Pierri[8,9], Gerarda Cappuccio[8,9], Irene Bruno[10], Lia Boyle[11], Vincenzo Nigro[10], Annalaura Torella[10], Tony Roscioli[12,13], Mark J Cowley[14,15,16], Sean Massey[1], Rhea Sonawane[17], Matthew D. Burton[18], Bitten Schonewolf-Greulich[19], Zeynep Tümer[19], Wendy Chung[20], Wendy A. Gold[21,22,23^], John Christodoulou[1,2,6,24^]

1 Detailed NGS and variant filtering method

Whole genome sequencing (WGS) was performed on case 1 and her parents at Kinghorn Centre for Clinical Genomics (Garvan Institute, Australia) using previously published methods (Heimer et al., 2016). The variant filtering was performed using Seave database; their in-house variant filtration platform (Gayevskiy, Roscioli, Dinger, & Cowley, 2019).

Agilent's ClearSeq Inherited Disease Panel containing 2742 genes was used to perform targeted sequencing on case 2 and 3 and their parents. The libraries were prepared using SureSelectQXT kit and samples were run using NextSeq500 sequencer at the NGS Facility (The Telethon Institute of Genetics and Medicine, Italy). Variant detection and filtering were done using their in-house tool based on GATK best practices (Musacchia et al., 2018).

Singleton targeted sequencing using Custom Early-Onset Epileptic Encephalopathy NGS panel including *MECP2* (https://www.fulgentgenetics.com/Epilepsy-Epileptic) was performed on case 4's genomic DNA at Fulgent Genetics facility. DNA libraries were prepared using hybrid capture technology and then sequenced using company's proprietary NGS technology. Variants were detected in genes with at least 10X coverage and curated manually using locus specific databases and literature searches. All variants with quality score less than 500 were confirmed using Sanger Sequencing.

2 Detailed case reports

Case 1 is a 19-year old female who is the first child of unrelated Vietnamese parents. She had static development from around 12 months and then was noted to have global development delay from around 15 months of age [crawling (20 months), standing (2 years), walking (5 years) and later stopped walking (7 years)]. She had developmental regression from 16 months with loss of communication skills and then loss of hand skills. She had seizures from 3 years of age treated with carbamazepine. Hence, the developmental regression pre-dated the onset of seizures. Other clinical features included classic handwringing stereotypies of RTT, episodic breath-holding, uncontrolled seizures (between 3-5 years), teeth grinding, small hands & feet and cool periphery, acquired microcephaly and poor eye contact. As per the revised Neul clinical diagnostic criteria she was diagnosed with classic RTT (Table 1) (Neul et al., 2010). Genetic screening for *MECP2*, Fragile X and Angelman Syndrome was negative. In addition, she has progressive lower limb spasticity, mild optic nerve atrophy and gait abnormalities (Table S2).

The brain CT scan performed in 2003 (13 years of age) showed loss of periventricular white matter and cerebral atrophy with ex vacuo dilatation of the lateral ventricles which was found to be more marked on the right-hand side. The first MRI was performed in 2004 and showed loss of volume of cerebral hemispheres (right more than the left) with prominence of the cortical sulci. Moreover, loss of white matter with small areas of gliosis was also observed in right deep white matter. Dilatation of the supra tentorial ventricular system with gross dilatation of the right lateral ventricle was also recorded. In addition, a small porencephalic cyst was also seen anterior to the body of the right lateral ventricle. Five years later (in 2009), another brain MRI was performed. There was global loss of brain volume which was most marked in the right frontoparietal and occipital lobes with associated ex vacuo dilatation of the right lateral ventricle. The persistent foci of signal abnormality, particularly within the deep white matter of the right cerebral hemisphere was still observed. The small pineal cyst remained unchanged. Also, reports showed that the tortuosity of the vessels in the region of the internal cerebral veins may be related to previous venous abnormality.

Case 2 is a 10-year old Italian female with, developmental delay, severe hypotonia, microcephaly, stereotypic hand movements, obesity, gait abnormalities, scoliosis and seizures. Development milestones were severely delayed but without regression; at the present time she

cannot maintain the sitting position without support, she cannot pull herself to stand, she can grasp large objects but she is not able to pincer grasp, and she is not verbal. She developed atonic seizures at 4 years of age and was started on clonazepam with partial control of the seizures. On brain MRI she was found to have cerebral and cerebellar atrophy, and optic nerve and chiasma atrophy. At 8 years of age, she did not fulfil Neul's clinical diagnostic criteria for RTT, but because of the significant clinical overlap we have classified her as RTT-like, based on two main and two supportive criteria (Table 1 and Table S2). She did not undergo newborn screening for medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, but had a biochemical diagnosis of medium-chain acyl-CoA dehydrogenase deficiency based on persistently elevated plasma C6, C8, C10 and C10:1 acylcarnitine levels as part of metabolic investigations. She was heterozygous for the common known c.985A>G (p.(Lys304Glu) pathogenic mutation in *ACADM* gene. A mutation in the second allele was not identified, and so far she has not had episodes of hypoglycemia or metabolic decompensation. So, whilst it is possible that she is merely a carrier of *ACADM* carrier, the persistent biochemical abnormalities mandate that she should be treated as if she has MCAD deficiency. Thus, she has been on carnitine therapy since the age of two years.

Case 3 is a 6-year old male from Italy with global developmental delay and intellectual disability. He smiled at 6months, rolled over at 8months, sat without support at 10months, and spoke his first words at 3 years of age. During last assessment at 4.6 years of age, it was found that he mostly crawls, but can walk if supported. He exhibited occasional stereotypic hand movements, abnormal muscle tone and inappropriate laughing/ screaming spells. Psychomotor delay was assessed using Bayley III scale at 4 years of age with the following scores; cognitive development IQ 42, receptive language 12 and expressive language 10, fine movements development 15, gross motor 35 and adaptive behavior 80. Repeated EEGs showed abnormal activity without seizures. He wore glasses for correction of astigmatism and hyperopia. He was started on growth hormone in March 2017 after being diagnosed with growth hormone deficiency and short stature. At that time, his height, weight, and head circumference were all below the 3%, though his height (94 cms), weight (12.1 kg) and head circumference (46.5cm) were proportional. He has been classified as RTT-like, as he fulfils two main and three supportive Neul's diagnostic criteria for RTT.

Case 4 (now deceased) was a 5.5 years old French-Canadian female with grossly abnormal psychomotor development since birth. At 6 months of age she lost the ability to breastfeed

followed by loss of ability to feed with bottles at 20 months and has required NG/G tube feeding since then due to repeated aspiration issues. She was never been able to verbalize and never had any pre-verbal vocalization.. The proband repeatedly mouth and chew her hands and never developed purposeful hand movements. She did not had a history of repetitive hand movements (wringing/clapping/rubbing) but exhibited, small, cold, hands and feet, history of bruxism, hypotonia and was non-ambulatory. She had a history of scoliosis and a high pain tolerance, and a diagnosis of epilepsy with her first seizure at one year old. She was on ketogenic diet for seizures and had seen some improvement/reduction in seizure activity. She had no history of peripheral vasomotor disturbance or inappropriate laughing/screaming spells, but periodical inconsolably crying spills with occasional sleep difficulties were reported. In addition, she had non-progressive cataracts, cortical visual impairment (diagnosed at 6 months of age) and optic nerve abnormalities. She was classified as RTT-like patient as she exhibits one main and five supportive criteria. *MECP2* testing was normal. Along with likely pathogenic *KIF1A* variant, a second variant was also identified in this patient in *GRIN2A* [NM_000833.4; c.4220C>T; NP_000824.1:p.(Ser1407Leu)] which was classified as variant of unknown significance and parental segregation was not performed.

3 Supplementary Figures

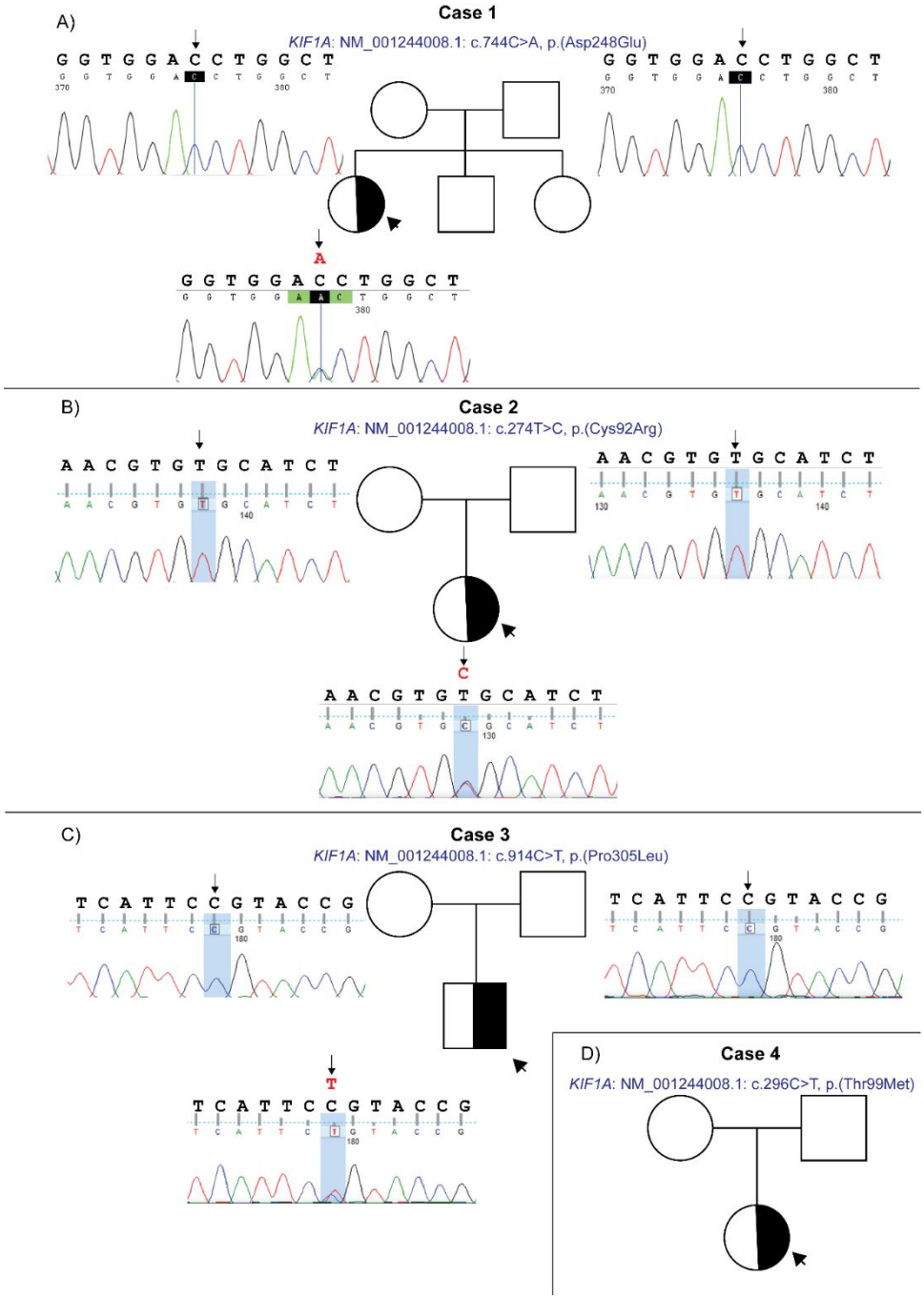


Figure S1: Pedigree of all four cases and Sanger sequence confirmation of *de novo* variant in in KIF1A.

A) Pedigree shows that case 1 is the affected female born to a non-consanguineous parent and has two healthy siblings. Sanger sequences show confirmation of *de novo* heterozygous variant *KIF1A*: NM_001244008.1:c.744C>A, p.(Asp248Glu). B) For case 2 and 3, pedigrees show that they are the only affected child born to non-consanguineous parents. Sanger sequences show confirmation of *de novo* heterozygous variant *KIF1A*: NM_001244008.1:c.274T>C, p.(Cys92Arg) and *KIF1A*: NM_001244008.1: c.914T>C, p.(Pro305Leu) respectively. D)

Pedigree for case 4 shows that she is the only affected child of non-consanguineous parents and Sanger sequencing was performed by Fulgent Genetics (USA).

```
C elegans      -MSSSVKVAVRVRPFNQREISNTSKCVLQVNGNTTTINGHSINKE----NFSFNFDHSYW 55
D melanogaster --MSSVKVAVRVRPFNSREIARESKCIIEMAGATTAITNPKVPPNTSDSVKRFNFDYSYW 58
H sapiens      MAGASVKVAVRVRPFNSREMSRDSKCIIQMSGSTTTIVNPKQPKET---PKSFSFDYSYW 57
M musculus     MAGASVKVAVRVRPFNSREMSRDSKCIIQMSGSTTTIVNPKQPKET---PKSFSFDYSYW 57
R norvegicus   MAGASVKVAVRVRPFNSREMSRDSKCIIQMSGSTTTIVNPKQPKET---PKSFSFDYSYW 57
                  :************.**::. ***:::: * **:* . .   :       *.**:***

C elegans      SFAR-NDPHFITQKQVYEELGVEMLEHAFEGYNVCIFAYGQTGSGKSYTMMGKANDPDEM 114
D melanogaster SHDH-HDADFSTQSMVYKDIGEEMLQHSFDGYNVCIFAYGQTGAGKSYTMMGRQE-EQQE 116
H sapiens      SHTSPEDINYASQKQVYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMMGKQE-KDQQ 116
M musculus     SHTSPEDINYASQKQVYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMMGKQE-KDQQ 116
R norvegicus   SHTSPEDINYASQKQVYRDIGEEMLQHAFEGYNVCIFAYGQTGAGKSYTMMGKQE-KDQQ 116
               *.   .* .: :*. **.::* ***:*:*:*************:********: :  ::

C elegans      GIIPRLCNDLFARIDNNNDKDVQYSVEVSYMEIYCERVKDLLNPNSGGNLRVREHPLLGP 174
D melanogaster GIIPMICKDLFTRIQDTETDDLKYSVEVSYMEIYCERVRDLLNPKNKGNLRVREHPLLGP 176
H sapiens      GIIPQLCEDLFSRINDTTNDNMSYSVEVSYMEIYCERVRDLLNPKNKGNLRVREHPLLGP 176
M musculus     GIIPQLCEDLFSRINDTTNDNMSYSVEVSYMEIYCERVRDLLNPKNKGNLRVREHPLLGP 176
R norvegicus   GIIPQLCEDLFSRINDTTNDNMSYSVEVSYMEIYCERVRDLLNPKNKGNLRVREHPLLGP 176
               **** :*:***:**::.  .::.***************:*****:. *************

C elegans      YVDDLTKMAVCSYHDICNLMDEGNKARTVAATNMNSTSSRSHAVFTIVLTQKRHCADSNL 234
D melanogaster YVEDLSKLAVTDYQDIHDLIDEGNKARTVAATNMNETSSRSHAVFTIFFTQRRHDLMTNL 236
H sapiens      YVEDLSKLAVTSYNDIQDLMDSGNKARTVAATNMNETSSRSHAVFNIIFTQKRHDAETNI 236
M musculus     YVEDLSKLAVTSYNDIQDLMDSGNKARTVAATNMNETSSRSHAVFNIIFTQKRHDAETNI 236
R norvegicus   YVEDLSKLAVTSYNDIQDLMDSGNKARTVAATNMNETSSRSHAVFNIIFTQKRHDAETNI 236
               **:**:*:** .*:** :*:*.*************.*********.*.:**:**   :*:

C elegans      DTEKHSKISLVDLAGSERANSTGAEGQRLKEGANINKSLTTLGLVISKLAEEST----KK 290
D melanogaster TTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEVASK---KK 293
H sapiens      TTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEMDSGPNKNK 296
M musculus     TTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEMDSGPNKNK 296
R norvegicus   TTEKVSKISLVDLAGSERADSTGAKGTRLKEGANINKSLTTLGKVISALAEMDSGPNKNK 296
               *** **************:****:* ***************** *** ***  :    :*

C elegans      KKSNKGVIPYRDSVLTWLLRENLGGNSKTAMLAALSPADINFDETLSTLRYADRAKQIVC 350
D melanogaster NTKKADFIPYRDSALTWLLRENLGGNSKTAMIAAISPADINYDETLSTLRYADRAKQIVC 353
H sapiens      KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAKQIRC 356
M musculus     KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAKQIRC 356
R norvegicus   KKKKTDFIPYRDSVLTWLLRENLGGNSRTAMVAALSPADINYDETLSTLRYADRAKQIRC 356
               :..: ..******.*************:***:**:******:*************** *

C elegans      QAVVN     355
D melanogaster KAVVN     358
H sapiens      NAVIN     361
M musculus     NAIIN     361
R norvegicus   NAIIN     361
               :*::*
```

Figure S2: Multiple sequence alignment showing the evolutionary conservation of KIF1A motor domain.

Sequences were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) by using RefSeq of *H sapiens* (NP_001230937.1), *M musculus* (NP_032466.2), *R norvegicus* (XP_017452404.1), *C elegans* (NP_001022041.2) and *D melanogaster* (NP_001246373.1).

* (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a . (period) indicates conservation between residues with weakly similar properties

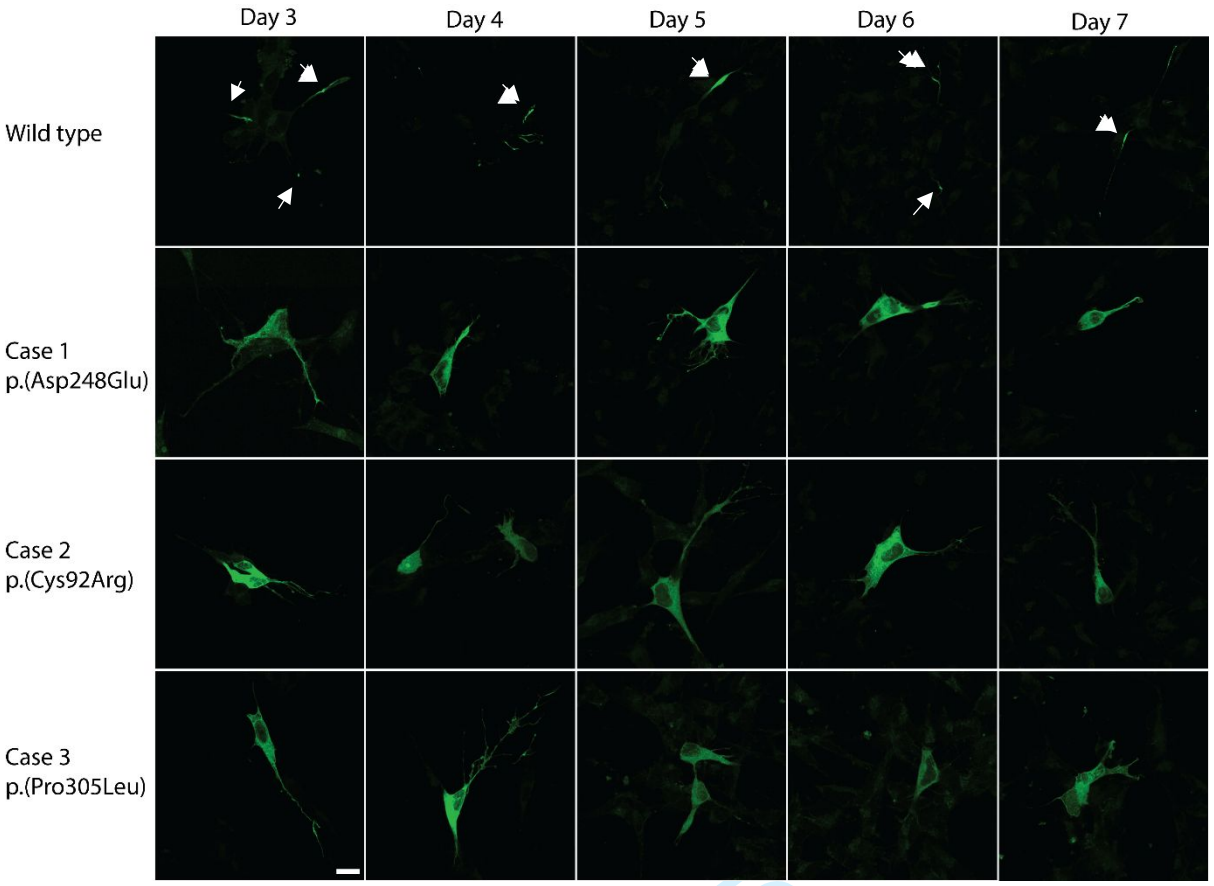


Figure S3: Time course evaluation of truncated KIF1A using the neurite tip accumulation assay in differentiating SH-SY5Y cells.

Representative images showing the localization pattern of rat KIF1A(1-393)-3xmCit (green) in differentiating SH-SY5Y cells over a seven day time course. The wild type KIF1A(1-393)-3xmCit motor protein was observed to accumulate in the distal tips (arrows) of differentiated SH-SY5Y cells as early as day 3 and was consistently observed in the neurite tips through day 7. In contrast, the mutant proteins were localised near the cell body across all time points. Scale bar = 10µm.

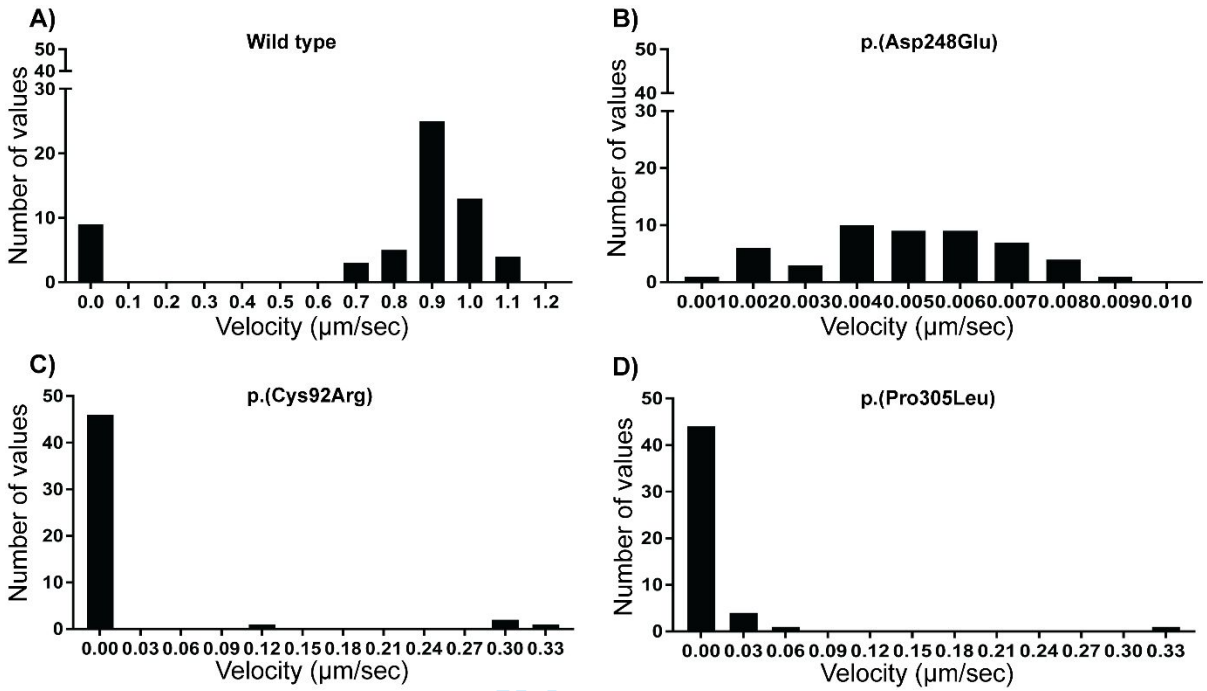


Figure S4: *In vitro* gliding assay experimental set up and frequency distribution of average velocity.

A) – D) Histograms of microtubule gliding velocities observed for the wildtype, p.(Asp248Glu), p.(Cys92Arg) and p.(Pro305Leu) proteins in the *in vitro* microtubule gliding assay. Average velocity (µm/sec) is on x-axis and number of values (observed events) are on y-axis. Motility events (n=50 for each experiment) were analyzed across three independent experiments. E) Schematic of the microtubule gliding assay in a flow chamber modified from (Yue et al., 2018). The biotinylated KIF1A(1-393)-LZ-mNG-AviTag motor proteins expressed in COS-7 cells were immobilized on a NeutrAvidin-coated glass coverslip. The motility mix containing rhodamine-labelled microtubules, ATP, and oxygen-scavenging system was then added into the flow chamber. Time lapse images were captured using the TIRF microscope.

4 Supplementary Tables

Table S1: List of primers used to introduce and confirm variants in *KIF1A*(1-393) construct

Primer name	Primer sequence (5' to 3')	Length	Tm (°C)	Purification
Site directed mutagenesis (SDM)				
Case 1_SDM_F	CAGCCTGGTGGAACTGGCGGGCAGT	25	78.38	PAGE
Case 1_SDM_R	ACTGCCCGCCAGTTCCACCAGGCTG	25	78.38	PAGE
Case 2_SDM_F	CCATAGGCAAAGATGCGCACGTTGTAGCCCTCA	33	79.14	PAGE
Case 2_SDM_R	TGAGGGCTACAACGTGCGCATCTTTGCCTATGG	33	79.14	PAGE
Case 3_SDM_F	ATACCGAGTCTCGGTAAAGGATGAAGTCTGTCTTC	35	78.1	PAGE
Case 3_SDM_R	GAAGACAGACTTCATCCTTTACCGAGACTCGGTAT	35	78.1	PAGE
Sanger confirmation				
Case 1_Sanger_F	CCTCCTGAACCCCAAGAACAA	21	59.86	Desalted
Case 1_Sanger_R	TTGCAGCGGATCTGTTTGGC	20	61.87	Desalted
Case 2_Sanger_F	AAATGGGCGGTAGGCGTG	18	60.44	Desalted
Case 2_Sanger_R	CAGGTTGCCTTTGTTCTTGGG	21	59.93	Desalted
Case 3_Sanger_F	CCTCCTGAACCCCAAGAACAA	21	59.86	Desalted
Case 3_Sanger_R	TTGCAGCGGATCTGTTTGGC	20	61.87	Desalted

Table S2: Evaluation of *KIF1A* associated features in our cases.

Category	Case 1	Case 2	Case 3	Case 4 (deceased)
Neurological Issues				
Hypotonia	No	Yes (Severe)	Yes (first 2 years of life)	Yes (since birth)
Hypertonia	Yes (lower limbs)		No	Yes (since birth)
Microcephaly	No	Yes	Yes	No
Global developmental delay	Yes	Yes	Yes†	Yes
Intellectual disability	Severe	Yes	Severe‡	Yes
Ataxia	Yes	Yes	No	NA
Spastic paraplegia	Yes (lower limbs)	Yes	No	NA
Peripheral neuropathy	No	Not evaluated	No	Yes
Nerve conduction studies	Not performed	Not performed	No	Yes; EMG at 10m$
Seizures	Yes (past history)	Yes	No¶	Yes
MRI changes				
Year (Age) when test was performed	2002 (2 year)	2013 (4 year)	2016 (3 year)	7 month
Cerebral atrophy	Yes	Yes	Yes!	No
Cerebellar atrophy	No	Yes	Yes	No
Reduced cerebral white matter	No	Yes	No	Reduced white matter: unclear if cerebellar
Thin or absent corpus callosum	No	Yes	No	No
Cerebellar abnormalities	No	Yes	No	No
Other abnormalities	small pineal cyst, tortuosity of the vessels (internal cerebral veins)	Yes	No	No
Muscle issues				
Upper limb spasticity	No	No	No	NA
lower limb spasticity	Yes (Progressive)	Yes	No	NA
Muscle hypertrophy	No	No	No	No
Muscle weakness	No	Yes	Yes	Yes
Gait abnormalities	Yes	Yes	Yes	Yes (non-ambulatory)
Ophthalmological Issues				
Cataracts	No	No	No	Yes
Cortical visual impairment	No	Yes	No	Yes
Optic nerve abnormalities	Yes (mild optic nerve atrophy)	Yes	No	Yes
Ptosis	No	No	No	No
Strabismus	No	Yes	No	No
Others				
Gastroesophageal reflux disease	No	No	No	Yes
Eczema	No	Skin xerosis on hands	No	No
Genital abnormalities	No	No	No	No

†: treated with growth hormone; ‡: global development score 45 (scale Bayley III), language 42, gross motor and fine 43; ¶: slow paroxysmal bilateral posterior activity in sleep; ! periventricular white matter (temporal

and occipital) alterations of signal; $: EMG at 10m showed 'anomaly in axon' and as per the mother of case 4: muscle biopsy at 14m showed "light axonal neuropathy with mitochondrial anomalies in the nerve"

5 Supplementary videos

Video 1: Example of recording of *in vitro* microtubule gliding assay for the wild type protein

Video 2: Example of recording of *in vitro* microtubule gliding assay for the p.(Asp248Glu) protein

Video 3: Example of recording of *in vitro* microtubule gliding assay for the p.(Cys92Arg) protein

Video 4: Example of recording of *in vitro* microtubule gliding assay for the p.(Pro305Leu) protein

6 Authorship contributions

S.K - design, investigation, analysis, Writing – Original Draft
N.V.B - Supervision, analysis, writing – review & editing
K.J.V - Resources (KIF1A constructs) and supervision on microtubule gliding assay methodology, writing – review & editing
C.J.N. - Supervision on total internal florescent microscopy, ImageJ software codes
B.B. – Assistance with optimization of microtubule gliding assay, writing – review & editing
Y.Y. – Assistance with optimization of microtubule gliding assay, writing – review & editing
C.E. - Clinician for Case 1
N.B.P - Clinician for Case 2, writing- review & editing
G.C. - Clinician for Case 2
I.B. - MRI report for Case 3
L.B. – Clinical information collected from family for Case 4
V.N. – Next generation sequencing for Case 2
A.T. – Next generation sequencing for Case 3
T.R. – Whole genome sequencing and data filtering for Case 1
M.J.C. – Developed bioinformatics pipeline for Case 1
S.M. – Assistance with experimental work

B.B. – Assistance with optimization of microtubule gliding assay, contributed figure 1C, writing – review & editing

Y.Y. – Assistance with optimization of microtubule gliding assay

R.S. – Placement student; helped with initial optimization of kymograph analysis

B.S. – Resources (collaboration with clinicians for Case 2 and 3)

Z.T. – Resources (collaboration with clinicians for Case 2 and 3), writing – review & editing

W.C – Clinical information collected from family for Case 4 – review & editing

W.A.G – Conceptualization, supervision, analysis, writing – review & editing

J.C. – Conceptualization, supervision, analysis, writing – review & editing, Lab head at MCRI

Author	ORCiD ID
SK	https://orcid.org/0000-0002-0689-2965
NVB	https://orcid.org/0000-0002-6768-3665
KJV	https://orcid.org/0000-0001-9329-4981
CN	https://orcid.org/0000-0002-8662-9840
CE	https://orcid.org/0000-0002-5752-5698
NBP	https://orcid.org/0000-0002-6895-8819
GC	https://orcid.org/0000-0003-3934-2342
IB	https://orcid.org/0000-0002-2437-6198
LB	https://orcid.org/0000-0002-8082-7125
VN	
AT	https://orcid.org/0000-0003-2479-6018
TR	https://orcid.org/0000-0003-1502-5000
MC	https://orcid.org/0000-0002-9519-5714
Y.Y.	https://orcid.org/0000-0002-8302-9778
B.B.	https://orcid.org/0000-0003-1212-3705
SM	https://orcid.org/0000-0002-0759-9306
RS	https://orcid.org/0000-0003-2541-3496
BSG	https://orcid.org/0000-0003-0088-0792
ZT	https://orcid.org/0000-0002-4777-5802
WC	https://orcid.org/0000-0003-3438-5685
WG	https://orcid.org/0000-0003-1808-0646
JC	https://orcid.org/0000-0002-8431-0641

Chapter 4

Chapter 4 Functional implications and phenotypic manifestations of *KIF1A* variants associated with other neurological disorders

4.1 Introduction

Traditionally, pathogenic variations in *KIF1A* have been largely associated with three major neurological disorders; autosomal recessive hereditary sensory neuropathy type IIC (HSNIIC), autosomal recessive hereditary spastic paraplegia 30 (SPG30) and autosomal dominant non-syndromic intellectual disability 9 (MRD9) (Erlich et al., 2011; Esmaeeli Nieh et al., 2015; Hotchkiss et al., 2016; Klebe et al., 2012; Krenn et al., 2017; Riviere et al., 2011). However, in 2017, natural history studies performed by our collaborators at Columbia University, USA (Dr. Wendy Chung and Ms Lia Boyle) suggested that different pathogenic autosomal dominant variations in *KIF1A* resulted in a more complex spectrum of neurological conditions, collectively referred to as KIF1A-associated neurological disorders (KANDs). The clinical phenotype of affected individuals, and the severity, depends upon a combination of several factors including position of the *KIF1A* variant and the inheritance pattern (L. Boyle & W. Chung, 2017; L. Boyle & W. K. Chung, 2017).

A subset of individuals with *KIF1A* variants develop HSNIIC, resulting in progressive degeneration of sensory neurons, progressive loss of sensation in hands and feet, as well as ulceration ultimately leads to chronic infection, and in some instances amputation of fingers and toes (Riviere et al., 2011). On the other hand, individuals with recessive SPG30 exhibit a slowly progressive spastic paraplegia mainly characterised by muscle weakness and stiffness, resulting in unsteady gait, and lower limb hyperreflexia (Cheon et al., 2017; Yoshikawa et al., 2019). The autosomal dominant form of KAND mainly results in a spectrum of clinical features that may vary in individuals affected with different *KIF1A* variants, but largely includes development delay, intellectual disability, speech delay, decreased muscle tone, muscle stiffness, muscle hypertrophy with age, scoliosis, microcephaly, optic nerve atrophy, ataxia, strabismus, ptosis, nystagmus, intention tremors and facial diplegia (L. Boyle & W. Chung,

2017). Although a natural history study has provided a greater understanding of the severity of clinical symptoms in individuals with KIF1A defects, however there is a huge gap in the evidence of pathogenicity of these variants and its correlation with the clinical features observed in individuals. In our KIF1A study cohort, all the 27 KAND individuals with 16 different missense variants exhibited a complex clinical phenotype. Although seven variants have been previously reported, to date only clinical and genetic reports are available with no further functional validation. Therefore, in this chapter the phenotype of all the 27 KAND individuals was compared and correlated with the functional consequences of variants affecting the KIF1A motor domain.

4.2 Results

4.2.1 Clinical summary of *KIF1A* individuals

Clinical features of 27 KAND individuals were compared as represented in Figure 4.1: Clinical features exhibited by 27 individuals with KAND. Detailed clinical information for all 27 individuals (case 5 to case 31) is available in Appendix E (Table E1). Overall, almost all the affected individuals exhibited global development delay (96%), intellectual disability (93%) and gait abnormalities (74%). Abnormal muscle tone was another common phenotype observed in these individuals, along with hypertonia (74%) or hypotonia (63%). Seizures were detected in only 37% of the affected individuals with *KIF1A* variants, however information was unavailable for a significant proportion of individuals (59%). A subset of individuals also exhibited microcephaly (33%), ataxia (33%) and peripheral neuropathy (26%). Of the 27 KAND individuals, spastic paraplegia was reported in 10 individuals (cases 6, 8, 17, 18, 26, 27, 28, 29, 30 and 31) and nerve conduction issues for only two individuals (cases 8 and 26), and so limited clinical information was available for a large proportion of individuals in this cohort. Several individuals also showed genital abnormalities such as a small penis and/or small scrotum (cases 7 and 10), enlarged clitoris (case 20) and small, retracted genitalia (case 26). Gastroesophageal reflux was reported in eight individuals (cases 7, 10, 11, 12, 14, 19, 24 and 25). KAND individuals in our cohort were not found to exhibit cataract and cerebral atrophy.

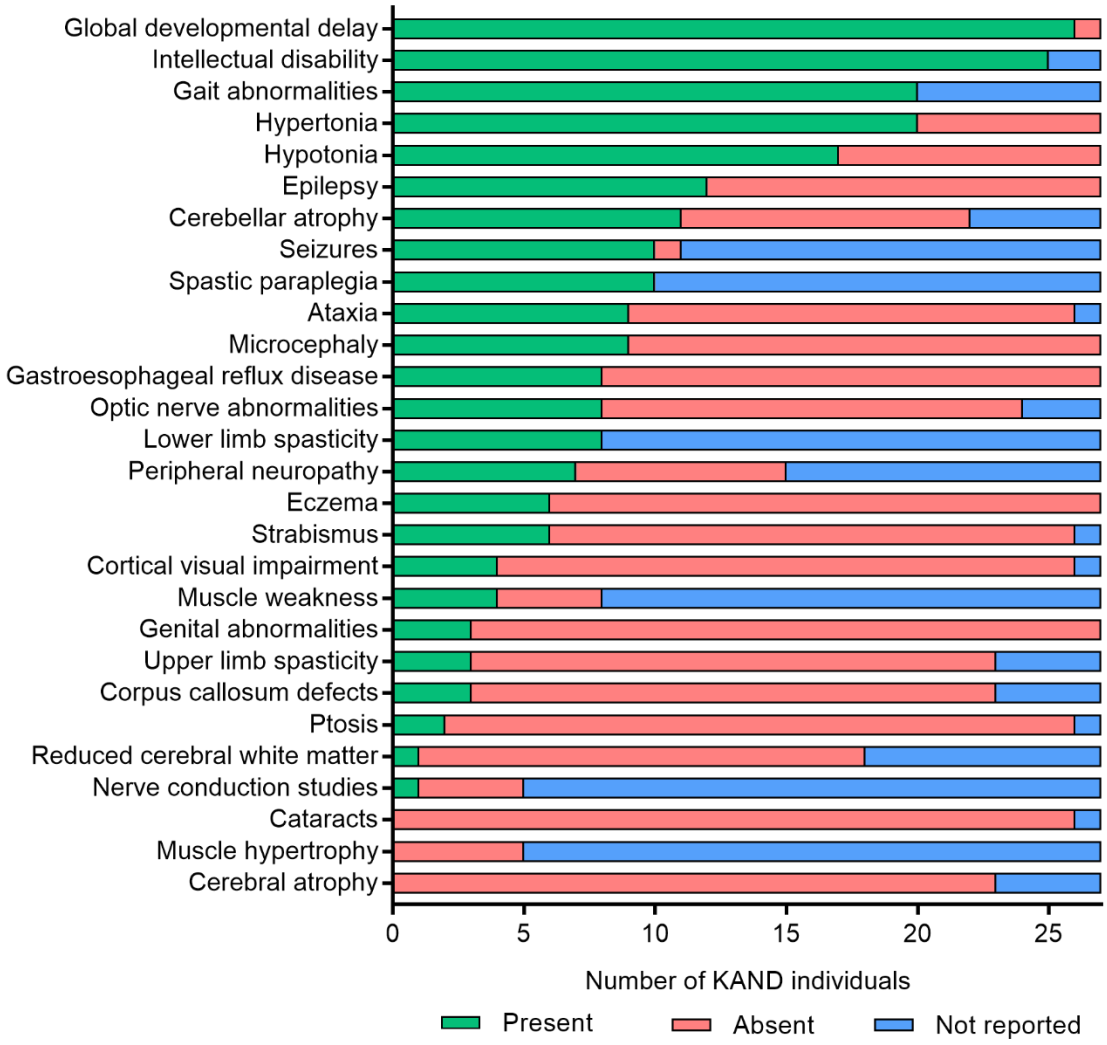


Figure 4.1: Clinical features exhibited by 27 individuals with KAND.

The majority of the individuals in KAND cohort exhibited development delay, intellectual disability and gait abnormalities. More than 50% of affected individuals also showed abnormal muscle tone. Seizures, microcephaly, spastic paraplegia and ataxia were present in about one third of the KAND individuals. A minor fraction of individuals exhibited microcephaly, optic nerve defects, spasticity, peripheral neuropathy, eczema. strabismus, cortical visual impairment, muscle weakness, genital abnormalities, ptosis and MRI abnormalities. Cataracts and cerebral atrophy were not observed as a common feature. Clinical information focused on seizures, spasticity, nerve conduction studies, muscle weakness and hypertrophy were not available for a large proportion of individuals in this cohort. Y-axis represents the clinical feature and x-axis represents the number of KAND individuals. Green = present, red = absent and blue = not reported.

The most common variants observed in our collected KAND cases were p.(Arg254Trp) and p.(Arg316Trp) affecting seven (cases 13 to 19) and five individuals (cases 26 to 30)

respectively. Of the five individuals with the p.(Arg316Trp) variant, two females were monozygotic twins of a triplet pregnancy.

Individuals with the p.(Arg254Trp) variant, along with development delay and intellectual disability, all exhibited hypertonia and hypotonia except case 19, where only hypertonia was observed with increased age. Seizures were observed in more than half of the individuals with p.(Arg254Trp) variant (cases 14, 16, 17 and 18). Only a few individuals exhibited microcephaly (cases 14 and 18) and peripheral neuropathy (cases 15 and 19). Cerebral atrophy was also a common feature among all individuals, except case 19. In addition, specific distinct neuro-imaging findings were observed including corpus callosum defects (case 15), reduced cerebral white matter (case 17) and a thin optic chiasm (case 19). Other individuals exhibited unique clinical features which were not observed in the rest of the individuals with the p.(Arg254Trp) variant. These less common features included optic nerve atrophy (case 18), strabismus (case 15), reflux and eczema (case 14) (Appendix E, Table E1).

All five individuals with the second most recurrent variant in our KAND cohort, p.(Arg316Trp), showed developmental delay and intellectual disability, except case 30, for whom no formal cognitive assessment was completed. Interestingly, all of the individuals exhibited spastic paraplegia and ataxia (except case 30 for whom information was not available). Peripheral neuropathy was only reported in one individual (case 26). Microcephaly (cases 29 and 30) and seizures (cases 26, 27 and 28) were only found to affect a subset of individuals with the p.(Arg316Trp) variant. Interestingly, of the individuals with the p.(Arg316Trp) variant, only case 26 had genital abnormalities, with small and retracted genitalia (Appendix E, Table E1). Although 23 year old monozygotic female twins (of triplets) showed many overlapping clinical features (Appendix E, Table E1), these females also exhibited several characteristic and distinct phenotypes, including hypertonia in case 28, and hypotonia in case 27. Moreover, brain MRI at 18 months of age in case 28 showed symmetrical cerebellar hypoplasia.

In summary, all the affected individuals with variants in the KIF1A motor domain demonstrated several overlapping but significant variable features that varied even among individuals with same amino acid substitution in KIF1A, further highlighting the complex nature of KANDs.

4.2.2 *In silico* analysis and variant classification

All the 16 *KIF1A* variants identified in 27 KAND individuals were confirmed to be *de novo* and heterozygous. Of these 16 variants, eight were novel, and eight were previously published. All these variants affected evolutionary conserved sites in KIF1A motor domain, which is a mutational hot spot region in KAND (Figure 4.2 and Table 4.1) (Gabrych, Lau, Niwa, & Silverman, 2019). All the variants were confirmed to be *de novo* by our collaborators. The *in silico* tools, MutationTaster and SIFT, predicted all these variants to be disease causing and likely damaging for KIF1A protein function. All these variants were absent in the population database, gnomAD, and were classified as likely pathogenic as per the ACMG guidelines (Richards et al., 2015). The Grantham score predicted nine variants [p.(Tyr89Asp), p.(Cys92Tyr), p.(Gly117Val), p.(Ser217Tyr), p.(Gly251Arg), p.(Arg254Trp). p.(Arg254Pro), p.(Arg307Pro), p.(Arg316Trp)] to cause a significant effect on protein structure (Grantham score >100) due to a major amino acid change; whereas the remaining seven variants [p.(Thr99Met), p.(Gly102Ser), p.(Asp248Gly), p.(Arg254Gln), p.(Leu278Pro), p.(Pro305Leu), p.(Thr344Met)] were predicted to have a minor effect, with a lower Grantham score (<100).

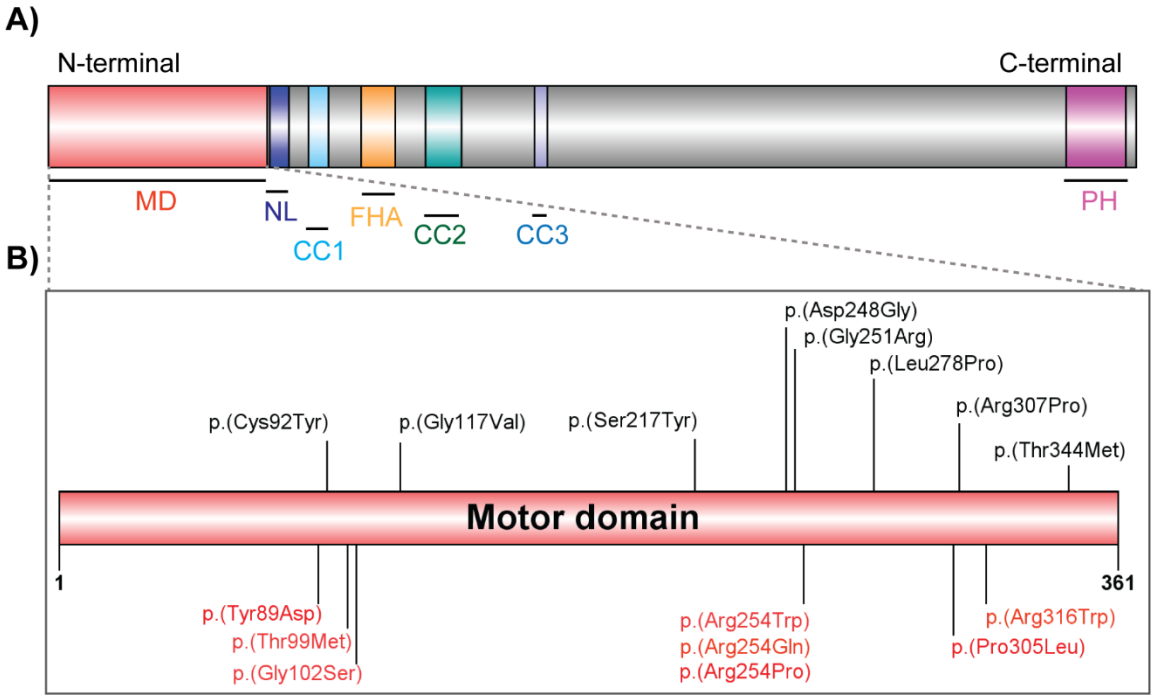


Figure 4.2 KIF1A domain structure and location of missense variant identified in KAND individuals

A) Full length KIF1A domain structure with various domains: motor domain (MD: 1-361 amino acid; red), neck coil (NC: 365-397 amino acid; blue), coiled coil domains [(CC1: 429-462 amino acid; cyan), (CC2; 622-681 amino acid; green) and (CC3: 801-822 amino acid; purple)], forkhead-associated domain (FH: 516-572 amino acid; orange), and pleckstrin homology domain (PH: 1676-1774 amino acid; pink). B) All the 16 individual KIF1A variants are clustered in the motor domain region, which is a reported variant hot spot in KIF1A (J. R. Lee et al., 2015). Eight novel variants are shown above the KIF1A motor domain structure (black text), while eight previously reported variants (red text) are represented below the protein domain structure.

Table 4.1 *In silico* prediction for *KIF1A* variants in individuals with KAND

	Case 5	**Case 6**	**Case 7**	**Case 8**	**Case 9**	**Case 10**	**Case 11**	**Case 12**
Variant	chr2:241727566 A>C; c.265T>G; p.(Tyr89Asp)	chr2:24172755 6C>T; c.275G>A; p.(Cys92Tyr)	chr2:24172753 5G>A; c.296C>T; p.(Thr99Met)	chr2:24172752 7C>T c.304G>A; p.(Gly102Ser)	chr2:24172748 1C>A; c.350G>T; p.(Gly117Val)	chr2:24172447 6G>T; c.650C>A; p.(Ser217Tyr)	chr2:24172321 1T>C; c.743A>G; p.(Asp248Gly)	chr2:24172320 3C>T c.751G>A; p.(Gly251Arg)
Novel or reported	Reported	Novel	Reported	Reported	Novel	Novel	Novel	Novel
Grantham score	160	194	81	56	109	144	94	125
	Cases 13-19	**Cases 20**	**Case 21**	**Case 22**	**Case 23**	**Cases 24, 25**	**Cases 26-30**	**Case 31**
Variant	chr2:241723194 G>A; c.760C>T; p.(Arg254Trp)	chr2:24172319 3C>T; c.761G>A; p.(Arg254Gln)	chr2:24172319 3C>G; c.761G>C; p.(Arg254Pro)	chr2:24172249 2A>G; c.833T>C; p.(Leu278Pro)	chr2:24171531 2G>A; c.914C>T; p.(Pro305Leu)	chr2:24171530 6C>G; c.920G>C; p.(Arg307Pro)	chr2:24171528 0G>A; c.946C>T; p.(Arg316Trp)	chr2:24171360 6G>A;c.1031C >T; p.(Thr344Met)
Novel or reported	Reported	Reported	Reported	Novel	Reported	Novel	Reported	Novel
Grantham score	101	43	103	98	98	103	101	81
Segregation	*de novo*							
Zygosity	heterozygous							
In silico tools	Disease causing [MutationTaster (score = 0.999)], damaging [SIFT]							
gnomAD	Absent							
ACMG	Likely pathogenic							
Location	Motor domain							

Genbank transcript ID: NM_001244008.1 (longest isoform): Position: hg19 chr2:241,653,181-241,759,725, size: 106,545 and total exon count: 49; NCBI Protein ID: NP_001230937.1 (longest isoform); Ensembl transcript ID: ENST00000320389. ACMG= American College of Medical Genetics, gnomAD = Genome Aggregation Database.

4.2.1 Three-dimensional (3D) protein modelling

In solution, full length KIF1A is present in an inactive dimeric state in the absence of cargo, known as auto-inhibition, which is proposed to occur to avoid consumption of ATP when cargo is unavailable for transportation (Hammond et al., 2009). The two motor domains of KIF1A alternately undergo ATP hydrolysis that results in the processive step-wise movement of KIF1A along the network of microtubules (Hirokawa et al., 2009; Stanhope & Ross, 2015). All of the 16 identified variants in KAND individuals were present in the motor domain of the KIF1A protein (Figure 4.2)

We used 3-dimensional modelling with HOPE program (Venselaar et al., 2010) to predict the likely conformational changes to KIF1A due to individual variants (Figure 4.3). Based on the crystal structure of the motor domain (amino acids 1 – 393), there are three main regions: Switch I (helix α3, loop L9, and β sheet β12), Switch II (L11, α4, L12, α5, and L13) and phosphate binding loop (P-loop) with a Walker fold ($GXXXXG^{102}K[T/S]$) that undergo conformational changes during ATP hydrolysis (Kull & Endow, 2013). The variant in case 11 [p.(Asp248Gly)] is present in Switch II and thus, is predicted to affect ATP binding. The negatively charged wild type (Asp) amino acid coordinates the metal ion (Mg^{2+}), critical for ATP binding, and the loss of negative change charge in the variant (Gly) residue is predicted to destabilize the interaction with the Mg^{2+} ion (Figure 4.3). Moreover, the wild type residue forms hydrogen bond with residues Ser104, a salt bridge with Lys103 and Arg203. However, the change in the size and hydrophobicity of the variant residue (Gly) is predicted to interfere with these interactions and disturb correct protein folding, thus, affecting the native KIF1A motor domain structure. The variants in case 7 [p.(Thr99Met)] and case 8 [p.(Gly102Ser)] were previously reported to affect the P-loop structure (nucleotide binding pocket) and disrupt ATP binding to the motor domain, thereby altering the ATP hydrolysis rate (J. R. Lee et al., 2015). The p.(Gly102Ser) variant specifically affects the conserved Walker fold ($GXXXXG^{102}K[T/S]$), causing a significant conformational change in the native structure, ultimately affecting ATP turnover (J. R. Lee et al., 2015).

When ATP is hydrolysed into ADP and γ –phosphate, the latter is released from the P-loop of KIF1A through a salt bridge structure formed by ionic interactions between residues Arg216 (loop L9) and Glu253 (loop L11). The variants in cases 13 to 21 [p.(Arg254Trp), p.(Arg254Glu) and p.(Arg254Pro)] are present next to the loop L11, and these were predicted to result in defective impaired γ –phosphate release and ATP binding (Ohba et al., 2015).

Along with ATP hydrolytic activity, efficient binding of the KIF1A motor domain to the microtubules is critical to prevent dissociation from the microtubules, and allow movement toward the distal ends of the neurite projections. Loops L8, L11 and L12 in the KIF1A motor domain interact with the microtubules, and this interaction is facilitated by the negative charge on the tubulin monomer in microtubules, which results in processive movement of KIF1A over long distances (Kikkawa et al., 2001; Nitta, Kikkawa, Okada, & Hirokawa, 2004). The Arg316 residue in the α5 helix (Switch II) further stabilizes the loop L8 for microtubule binding. The variant in cases 26-30 [p.(Arg316Trp)] was reported to disrupt the stable interaction between the motor domain of KIF1A and negatively charged tubulin, thereby compromising the interaction between KIF1A motor domain and microtubule (J. R. Lee et al., 2015).

Perturbed local conformation of the protein core exerts significant changes in the overall structure and function of the highly conserved motor domain of the KIF1A protein. The wild type Proline (Pro) in case 23 [p.(Pro305Leu)] was comparatively smaller, buried in the core and provided more conformational rigidity to the KIF1A motor domain structure, and so the variant Leucine (Leu) was predicted to affect the correct folding of the protein. The small, neutral, hydrophobic wild type residue Tyrosine (Tyr) has been replaced with the comparatively larger, negative Aspartic acid (Asp) at position 89 in case 5 [(p.(Tyr89Asp)] and was predicted to impair the interactions between the variant residue and other neighbouring residues. The variant hydrophobic residue Valine (Val) in case 9 [p.(Gly117Val)] and positively charged Arginine (Arg) in case 12 are larger and less flexible compared to the wild type residue Glycine (Gly). These substitutions were predicted to alter the torsion angles that could force the local backbone into an incorrect conformation. The wild type residue Serine (Ser) and Cysteine (Cys) are buried in the core of the protein and forms hydrogen and disulphide bonds respectively. However, in case 10 [(p.(Ser217Tyr)] and case 6

[p.(Cys92Tyr)], the variant residue Tyrosine (Tyr) is larger in size and is predicted to not fit correctly, thereby affecting the correct folding of KIF1A motor domain. Similarly, the larger positively charged variant residue (Arg) in case 12 [p.(Gly251Arg)] replaces a smaller, neutral and hydrophobic wild type (Gly) amino acid buried in the core, and is predicted to alter correct protein folding. Additionally, in case 22 [p.(Leu278Pro)], the α-helix structure in the core of KIF1A motor domain is predicted to become destabilized in the absence of wild type Proline (Pro) residue, resulting in severe effects on the overall structure of the protein. The variant p.(Arg307Pro) in cases 24 and 25 was also predicted to affect protein-protein interactions, as the amino-acid substitution abolished a positive charge critical for ionic interactions. The variant replaced the wild type (Arg) residue with a small, neutral and more hydrophobic variant amino acid (Pro) that may result in loss of hydrogen bonding with other residues, thereby disturbing correct protein folding. Similarly, the variant residue Methionine (Met) in case 31 [p.(Thr344Met)] is smaller and more hydrophobic than the wild type amino acid Threonine (Thr), which is buried in the core of the motor domain, and was predicted to result in loss of hydrogen bonding, resulting in misfolding of the critical motor domain. Overall, the three-dimensional modelling, revealed that all identified variants in KAND individuals were predicted to have severe effects on the structure of KIF1A protein, ultimately affecting the function of KIF1A.

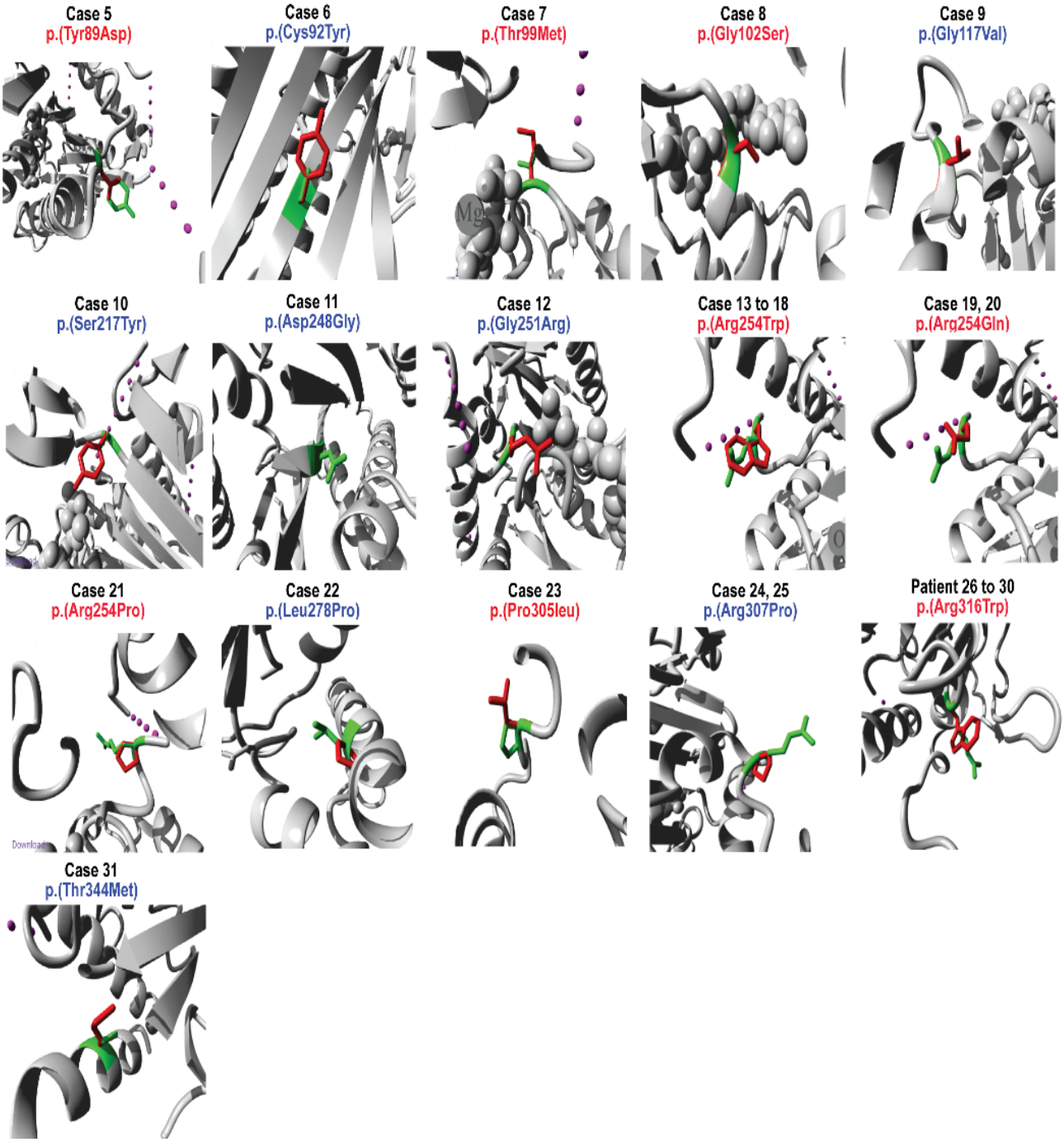


Figure 4.3 Molecular modelling of KIF1A variants in individuals with KAND

Schematic structures of the wildtype and variant amino acids with a close up of the variant site modelled with the HOPE resource (https://www3.cmbi.umcn.nl/hope). The protein is coloured in grey, the side chain of the wild type is green and variant is red. Previously published variants are labelled with red text whereas the novel variants are highlighted in blue text.

4.2.2 Neurite tip accumulation assay in differentiated SHSY5Y cells

Of the 16 variants identified in our 27 KAND individuals, eight variants were novel and eight variants were previously reported (Esmaeeli Nieh et al., 2015; Landrum et al., 2018; Pennings et al., 2020; Spagnoli, Rizzi, Salerno, Frattini, & Fusco, 2019) (Figure 4.2 KIF1A domain structure and location of missense variant identified in KAND individuals). Of these eight previously reported variants, the pathogenicity of only the p.(Thr99Met) variant has been previously reported, demonstrating the reduced accumulation of KIF1A motor domain protein harbouring the p.(Thr99Met) variant in the distal neurites of rat hippocampal neurons (Hamdan et al., 2011; J. R. Lee et al., 2015). However, for the remaining six previously reported variants [p.(Tyr89Asp), p.(Gly102Ser), p.(Arg254Trp), p.(Arg254Gln), p.(Arg254Pro) and p.(Arg316Trp)] identified in our 27 KAND individuals, only genetic findings with predictive *in silico* analysis have been reported. Thus, no functional validation of pathogenicity exists in the literature that relates the effect of each of these variants to the overall function of KIF1A.

We therefore assessed the consequences of these KIF1A variants using the neurite tip accumulation assay (as described in Chapter 2 and Chapter 3), a widely used method to assess the efficacy of kinesin movement along microtubules in various cell-based models. The full length KIF1A is inactive in solution due to an "auto-inhibition" regulatory mechanism. The two regions in KIF1A that contribute to this mechanism are the FHA and CC2 domains (amino acids 492–726) inhibiting microtubule binding and the CC1 domain (amino acids 394–491) preventing processive motility along microtubules (Hammond et al., 2009).Thus, for this assay, a KIF1A(1-393)-3xmCit construct containing the KIF1A protein sequence corresponding to amino acids 1 to 393 was tagged with three tandem copies of fluorescent mCitrine (mCit), a generous gift of Kristen Verhey (University of Michigan, Ann Arbor, MI). The KIF1A(1-393) sequence contained the motor domain (amino acid 1-361), neck linker, and neck coil domains of KIF1A, which are dimerized via a GCN4 leucine zipper (LZ) (Hammond et al., 2009). Individual-specific variants were introduced using site-directed mutagenesis into the KIF1A(1-393)-3xmCit construct. We determined the effect of identified variants on the ability of KIF1A to accumulate at the tip of the neurites in differentiated SH-SY5Y cells.

Differentiated SH-SY5Y cells transfected with the wild-type KIF1A(1-393)-3xmCit construct, showed a clear accumulation of KIF1A protein at the distal ends of the neurite projections, consistent with previous reports (Hammond et al., 2009; Huang & Banker, 2012). However, in case of the individual variants, a substantial accumulation of KIF1A protein remained localized in the cell body, close to the protein synthesis site. This demonstrates a severe effect on the capacity of most of the mutant variants of the KIF1A motor domain to move along microtubule network progressively. The mean fluorescence intensity (MFI) ratio of KIF1A(1-393)-3xmCit fluorescence along the neurite length versus the cell body was found to be significantly higher in the cells transfected with the wild-type motor protein compared to any of the individual variants (Figure 4.4). Of interest, case 7 individual variant p.(Thr99Met), although showing a comparably low MFI to that of the wild type, did not show a statistically significant difference (case 7). Interestingly, comparable MFI changes were observed in nine individuals (cases 13 to 21) harbouring likely pathogenic variants [(p.(Arg254Trp), p.(Arg254Gln) and p.(Arg254Pro)] affecting the same codon at residue Arg254 (Figure 4.4).

Thus all individual variants studied, except p.(Thr99Met), showed a statistically significant reduction in the of KIF1A accumulation at the cell periphery.

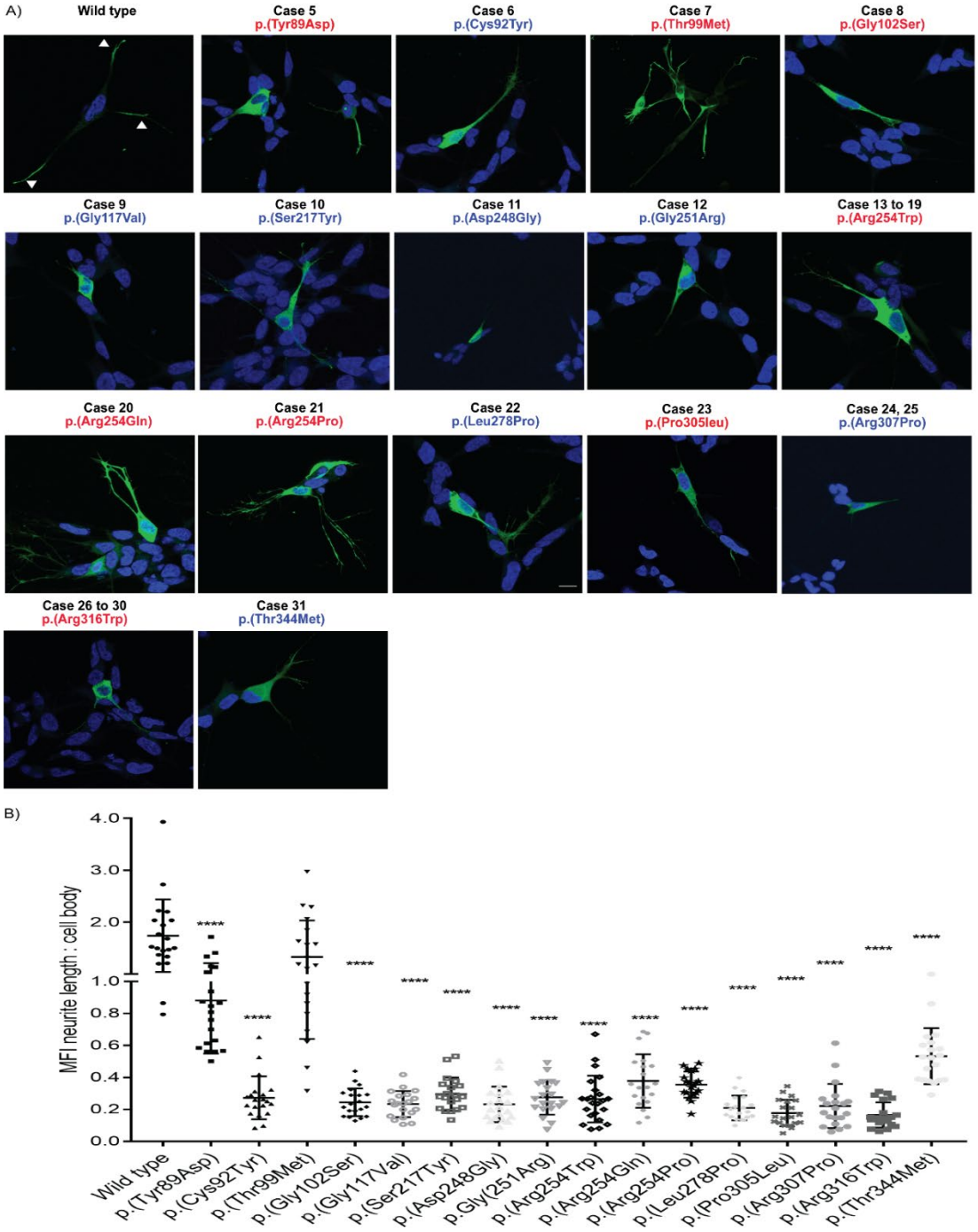


Figure 4.4: Neurite tip accumulation in differentiated SH-SY5Y cells for KIF1A variants in KAND individuals

A) Rat KIF1A (1-393), that is 100% homologous to human sequence, tagged with 3xmCit tag (green) accumulated in the distal tips of differentiated SH-SY5Y cells in case of wild-type, but was localised closer to the cell body in cells transfected with KIF1A constructs containing individual variants. B) A significant reduction in the mean florescence intensity (MFI) along the neurite length versus the cell body was observed in the 16 individual variants except p.(Thr99Met). Data was analysed from three independent experiments using Mann-Whitney test (n=20 from each experiment) and is plotted as mean ± standard deviation. **** represents $p<0.0001$ versus wild-type.

4.2.3 *In-vitro* microtubule gliding assay

Next, we wanted to determine whether the variants affected the movement of KIF1A along synthetic microtubules. To do this, we had established an *in-vitro* microtubule gliding assay, as described in Chapter 2, which mimics the movement of KIF1A along the microtubule network in the cell, but in an *in vitro* environment (Yue et al., 2018).

We performed this assay to demonstrate proof of concept using one of the most common variants [p.(Arg316Trp)] reported in the literature that affected five individuals in our KAND cohort. In this assay, wild type or p.(Arg316Trp) variant KIF1A(1-393)-LZ-mNG-AviTag construct was compared where C-terminal mNeonGreen (mNG; Allele Biotech) was a fluorescent tag, and AviTag corresponded to amino acid sequence (GLNDIFEAQKIEWHE) that was biotinylated by co-expressing HA-tagged bacterial biotin ligase BirA (Yue et al., 2018). Briefly, biotinylated KIF1A(1-393) motor protein prepared in COS-7 cells was first immobilized on a glass surface in a 10 µl flow chamber using a layer of NeutraAvidin protein before adding the motility mixture, containing fluorescent rhodamine labelled synthetic microtubules, ATP, $MgCl_2$ and oxygen scavenging solution. The ability of the KIF1A(1-393) motor protein to work together to move the microtubules along the surface was examined using fluorescence based total internal reflection microscopy (TIRF) that exclusively captures the signal from microtubules adjacent to the immobilized KIF1A protein, thus eliminating non-specific background from floating microtubules in the vicinity (Appendix C, Video C5).

From our results, we observed that a major fraction of the microtubules (~90%) were able to bind to the immobilized wild-type KIF1A(1-393) protein and moved with an average velocity of 0.91±0.09 µm/sec. However, the majority of the microtubules with the variant p.(Arg316Trp) protein could not bind and only ~10% of the bound microtubules were observed to move, with a statistically significant low average velocity of 0.0006986±0.002 µm/sec (Figure 4.5)

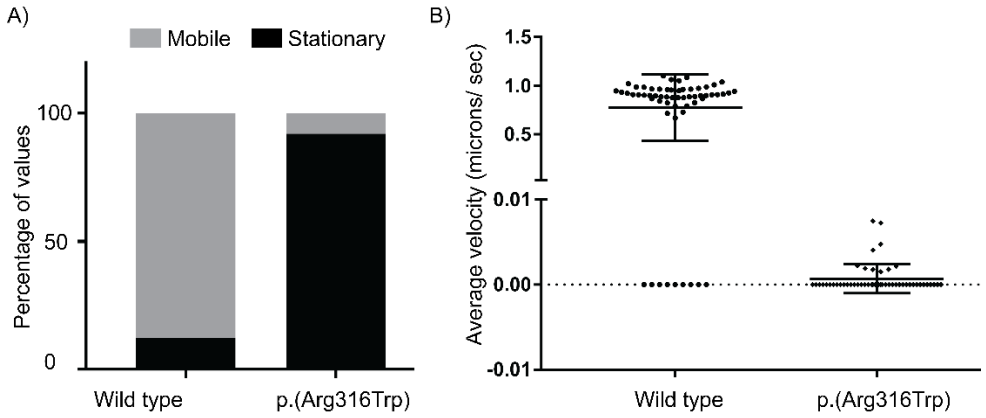


Figure 4.5: *in vitro* microtubule gliding for wild type and variant p.(Arg316Trp) motors.

A) Percentage of microtubules that were stationary (black) versus motile (grey). Approximately 90% of the bound microtubules moved on the surface of the wild type protein, whereas only 10% of the microtubules were able to glide on the surface of variant p.(Arg316Trp) protein.

In summary, these findings suggest that the variant p.(Arg316Trp) adversely affects the binding and movement of KIF1A motor protein along the cellular microtubule network.

4.3 Discussion

Previously published individuals with autosomal dominant KAND are known to exhibit a spectrum of clinical features including delay in attainment of developmental milestones, intellectual disability, peripheral neuropathy, ataxia, ptosis, scoliosis, abnormal muscle tone, muscle weakness and stiffness ultimately progressing to spastic paraplegia in a subset of individuals with KIF1A defects (L. Boyle & W. Chung, 2017). In addition, a significant number of previously published individuals have been reported to experience ocular abnormalities such as optic nerve atrophy, strabismus, cortical blindness, astigmatism, impaired visual acuity, involuntary eye movements, and partial ptosis (Esmaeeli Nieh et al., 2015; Ohba et al., 2015; Yoshikawa et al., 2019). Moreover, neuro-imaging studies have highlighted the occurrence of cerebellar atrophy, corpus callosum abnormalities and cerebral atrophy in KAND individuals. A variety of seizure types, including absence seizures, atonic drop seizures, focal seizures and complex seizures, have also been reported in a few individuals. Interestingly, abnormal EEGs have been reported in a subset of individuals without any overt seizure episodes (L. Boyle & W. Chung, 2017; Erlich et al., 2011; Esmaeeli Nieh et

al., 2015; Hotchkiss et al., 2016; Klebe et al., 2012; Krenn et al., 2017; Riviere et al., 2011). Other features that evolve in KAND individuals include difficulty with swallowing, gastrointestinal reflux, urogenital anomalies, kidney dysfunction, and eczema. Of note, no individual with KAND has been reported to exhibit any cardiac abnormalities, hearing impairment, auto-immune abnormalities or increased susceptibility to cancer (L. Boyle & W. Chung, 2017).

This is consistent with our KAND cohort, with almost all of 27 individuals exhibiting global developmental delay or intellectual disability except for the individual with the p.(Arg254Gln) KIF1A variant (case 20), although for two individuals (cases 6 and 30) no formal assessment was available. Other common features among all the studied KAND individuals included abnormal muscle tone, seizures, microcephaly and an abnormal gait. In addition, a few individuals had gastroesophageal reflux, eczema and genital abnormalities, features that now being recognised to be associated with KAND individuals (Erlich et al., 2011; Esmaeeli Nieh et al., 2015; Hotchkiss et al., 2016; Klebe et al., 2012; Krenn et al., 2017; Riviere et al., 2011).

All the identified variants in our KAND cohort were found to be *de novo*, heterozygous and clustered within the mutational hot spot region of the motor domain. This domain is critical for normal function as it is responsible for dimer formation and binding with ATP and microtubules. This enables KIF1A to move along the microtubules in step-wise manner using energy from ATP hydrolysis to carry specific cargo towards the end of the projections in neurons. Due to the paucity of studies correlating the effects of the variants to the overall function of KIF1A, we conducted two highly relevant functional assays, the neurite tip accumulation assay and the *in-vitro* microtubule gliding assay, to determine the pathogenicity of each variant. Almost all of the individual variants studied predicted to affect the protein folding showed significantly reduced accumulation of variant KIF1A motor protein at the tip of the neurites in the neurite tip accumulation assay.

The neurite tip accumulation assay showed similar defects in KIF1A movement for the the recurring likely pathogenic variants in our KAND cohort affecting residue Arg254

[(p.(Arg254GTrp), p.(Arg254Gln) and p.(Arg254Pro)] and Arg316 [p.(Arg316Trp)]. Variants in these individuals resulted in the change of positively charged wild type Arg residue to highly hydrophobic variant Trp residue. Apart from global developmental delay and intellectual disability, all these individuals exhibited distinct specific features that were not found in individuals affected with the same change in the KIF1A motor domain, consistent with the previously published findings (L. Boyle & W. Chung, 2017). Although hypertonia was observed in all three individuals, hypotonia was only evident in case 21. Interestingly, no global development delay was observed in case 20. A range of MRI abnormalities were observed in these individuals included an optic chiasm abnormality (case 19) and a small cisterna magna (case 21). Optic nerve abnormalities, which are another common feature of KAND, were only observed in case 20. Optic nerve pallor (end stage of optic nerve damage) was noted at 5 years of age without associated vision impairment. However, recent Optical Coherence Tomography (at 11 years of age) was found to be marginally abnormal, and the individual has had recent complaints of vision loss. Gastroesophageal reflux and an enlarged clitoris were only observed in case 19, and eczema was the feature seen only in case 21. The two common clinical features exhibited by all of these individuals included gait abnormalities, intellectual disability.

The p.(Arg316Trp) has been reported as a *de novo* heterozygous variant in a number of unrelated individuals with intellectual disability, spastic paraparesis, cerebellar atrophy and visual disturbances (Esmaeeli Nieh et al., 2015; J. R. Lee et al., 2015; Ohba et al., 2015). In our KAND cohort of individuals, p.(Arg316Trp) was found to be the second most common variant, occurring in five individuals (cases 26 to 30). Interestingly, although a few key features such as development delay, intellectual disability, nerve conduction abnormalities and gait abnormalities were consistently observed among all five individuals, there was variability in abnormal muscle tone, seizures, microcephaly, MRI abnormalities and optic nerve abnormalities was observed. Moreover, all five individuals with the p.(Arg316Trp) variant had spastic paraplegia consistent with previous reports (J. R. Lee et al., 2015; Ohba et al., 2015; Ylikallio et al., 2015). Of note, only case 26 was shown to have small, retracted genitalia. Interestingly, of these five cases, case 27 and case 28 were monozygotic twin females of triplets who showed a similar phenotype except for muscle tone, where hypertonia was observed in case 27, while case 28 had hypotonia. Molecular modelling predicted this variant [p.(Arg316Trp)] to cause reduced binding of KIF1A motor domain with the microtubule

network (J. R. Lee et al., 2015). Similar findings were observed for the *in-vitro* microtubule gliding assay, where a significant fraction of synthetic fluorescent microtubules could not bind to the variant protein, thus providing additional functional evidence that perturbation at residue Arg316 may result in reduced binding between KIF1A motor domain and microtubules.

The p.(Thr99Met) variant is a common recurrent variant found in published literature, and has been reported as a *de novo* heterozygous variant in six unrelated individuals with complex neurological phenotypes characterized by moderate to severe development delay and/or intellectual disability, epilepsy, spastic paraparesis, variable cerebellar atrophy, peripheral neuropathy, ataxia, visual disturbances and/or progressive neurodegeneration (Esmaeeli Nieh et al., 2015; Hamdan et al., 2011; J. R. Lee et al., 2015). Moreover, a recent report also found this variant in a female diagnosed with Progressive encephalopathy with Edema, Hypsarrhythmia and Optic atrophy (PEHO) syndrome which is a neurodevelopmental disorder characterized by infantile-onset progressive encephalopathy, intellectual disability, hypotonia, seizures, visual impairment, microcephaly and facial dysmorphism, further expanding the phenotypic spectrum of KAND (Langlois et al., 2016). Our case 7 with the p.(Thr99Met) variant had an overlapping phenotype with reported individuals including hypotonia, microcephaly, development delay, seizures, optic nerve hypoplasia and corpus callosum abnormalities. Interestingly, our case also exhibited strabismus and was non-ambulatory and had short stature. Other rare features seen in KAND individuals were also observed in case 7, including gastroesophageal reflux disease, eczema and a small scrotum. Functional consequences of this variant have been previously reported where the KIF1A motor domain harbouring the p.(Thr99Met) variant was found to have defective mobility, with reduced accumulation of the expressed variant protein observed in the distal neurites of rat hippocampal neurons (Hamdan et al., 2011; J. R. Lee et al., 2015). However, results from our neurite tip accumulation assay did not show statistically significant differences. This discrepancy could be a consequence of the large variation in MFI that we observed in these experiments.

In conclusion, through this study we have evaluated the functional consequences of the 16 different *KIF1A* variants affecting 27 KAND individuals exhibiting variable phenotypes. This study further establishes that different variants in the KIF1A motor domain can result in overlapping clinical features. We suggest that the specific features exhibited by the affected

individuals and the phenotypic severity still largely depend upon the variant, but it is likely that other additional factors contribute to disease severity that need to be further explored.

Chapter 5

Title

Expanding the genetic landscape of Rett syndrome to include Lysine (K) acetyltransferase 6A (*KAT6A*)

Authors

Simranpreet Kaur [a,b], Nicole J. Van Bergen [a,b], Bruria Ben-Zeev [c,d,e], Emanuela Leonardi [f,g], Tiong Y. Tan [b,h], David Coman [i,j,k], Benjamin Kamien [l,m], Susan M. White [b,h], Miya St John [n,o], Dean Phelan [h], Kristin Rigbye [h], Sze Chern Lim [h], Michelle C. Torres [h], Melanie Marty [h], Elena Savva [h], Teresa Zhao [h], Sean Massey [a], Alessandra Murgia [f,g], Wendy A. Gold [p,q,r], John Christodoulou [a,b,h,p]

Affiliations

a. Brain and Mitochondrial Research Group, Murdoch Children's Research Institute, Royal Children's Hospital, Melbourne, Australia.
b. Department of Paediatrics, University of Melbourne, Melbourne, Australia
c. Paediatric Neurology Institute, The Edmond and Lily Safra Children's Hospital, Israel
d. Chaim Sheba Medical Center, Tel HaShomer, Israel
e. Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
f. Molecular Genetics of Neurodevelopment, Department of Woman and Child Health, University of Padova, Italy
g. Fondazione Istituto di Ricerca Pediatrica (IRP), Città della Speranza, Padova, Italy
h. Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Melbourne, Australia
i. Department of Paediatrics, The Wesley Hospital, Brisbane, Brisbane, Australia
j. Queensland Children's Hospital, Brisbane, Australia
k. School of Medicine, University of Queensland, Brisbane, Australia
l. Genetic Services of Western Australia, Western Australia, Australia
m. Faculty of Health and Medical Sciences, University of Western Australia, Australia

n. Speech and Language, Murdoch Children's Research Institute, Parkville, Victoria, Australia
o. Department of Audiology and Speech Pathology, University of Melbourne, Parkville, Victoria, Australia
p. The University of Sydney, School of Medical Sciences and Discipline of Child and Adolescent Health, Faculty of Medicine and Health, Sydney, Australia
q. Molecular Neurobiology Lab, Kids Research, Westmead Children's Hospital, Westmead, Sydney, Australia
r. Kids Neuroscience Centre, Kids Research, Children's Hospital at Westmead, Westmead, Sydney, Australia

Corresponding author

Prof John Christodoulou

Brain and Mitochondrial Research Group

Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Victoria 3052 Australia, Email: john.christodoulou@mcri.edu.au

Manuscript information

Number of text pages (including references and figure legends): 22

Number of figures: two in main text and two in supplementary material

Word and character count

Number of words in abstract: 220

The total number of characters in the paper: Introduction, Results, Discussion, Materials and Methods, Acknowledgement = 23,612 (no spaces) and 27,502 (with spaces).

Abbreviation footnote

ACMG: American College of Medical Genetics

ASD: Autism Spectrum Disorder

BRPF1: Bromodomain And PHD Finger Containing 1

CDKL5: Cyclin dependent kinase like 5

EAF6: Esa1p-Associated Factor 6

FOXG1: Forkhead box G1

gnomAD: Genome Aggregation Database

HAT: Histone acetyltransferases

HDAC: Histone deacetylases

IGV: Integrated Genomic Viewer

KAT6A: Lysine (K) acetyltransferase 6A

KAT6B: Lysine (K) acetyltransferase 6B

MECP2/ MeCP2: Methyl-CpG-binding protein 2

MYST: MOZ, Ybf2/Sas3, Sas2, and Tip60

NEMM: N-term region of ENOK, MOZ or MORF

NGS: Next generation sequencing

PHD: plant homeodomain

PFO: Patent Foramen Ovale

PDA: Patent Ductus Arteriosus

PHD: Plant homeodomain

RTT: Rett syndrome

RUNX: Runt-related transcription factor

Ser/Met: serine- and methionine-rich transactivation domain

Abstract

The histone acetyltransferase family of proteins are involved in transcriptional regulation through chromatin remodelling. Next generation sequencing and functional validation studies have uncovered chromatin dysfunction as one of the most affected pathways associated with neurodevelopmental disorders such as Rett syndrome (RTT). RTT is a progressive, severe paediatric-onset neurological condition in which affected individuals may be classified as classic or atypical based on specific clinical criteria. Individuals who have some but not the requisite number of these criteria may be referred to as RTT-like. Although RTT is mainly caused by pathogenic variants in the X-linked Methyl-CpG-binding protein 2 (*MECP2*) gene that codes for a transcriptional regulator, up to 17% of individuals do not have a genetic diagnosis. In this study, we have identified seven individuals with RTT/RTT-like features harbouring variants in lysine (K) acetyltransferase 6A (*KAT6A*) that result in protein truncation and loss of the well conserved C-terminal domain critical for binding of transcriptional co-regulators. All the variants are clustered in exon 17 of *KAT6A* and are predicted to be pathogenic based on the ACMG criteria. Along with the RTT-specific symptoms, some individuals exhibit other clinical features that are often seen in affected children with *KAT6A*-related intellectual disability, such as facial dysmorphism, gastrointestinal problems and cardiac defects. This is the first report linking RTT/RTT-like clinical phenotypes with pathogenic variants in *KAT6A*.

Keywords

Rett syndrome; RTT-like; *MECP2*; *KAT6A;* chromatin regulator; KAT6A syndrome, *KAT6A*-related intellectual disability

1. Introduction

Pathogenic variants in Methyl CpG Protein 2 (*MECP2*; OMIM 300005) result in an X-linked, severe and progressive epigenetic disorder, Rett syndrome (RTT, OMIM: 312750) that predominantly affects females (Hagberg, Aicardi, Dias, & Ramos, 1983; Rett, 1966). The worldwide prevalence of RTT ranges from 1:10,000 to 1:23,000 live female births and is

considered the most common cause of severe intellectual disability in females after Down syndrome (Armstrong, Hangauer, Agazzi, Nunez, & Gieron-Korthals, 2010). Using Neul's revised diagnostic criteria, affected individuals can be clinically classified as classic or atypical RTT (Neul et al., 2010). After 6 to 18 months of apparently normal development, girls with the classic form of RTT gradually start losing their previously acquired skills including language, communication, co-ordination and social interaction. Along with acquired microcephaly, affected girls lose purposeful use of their hands and develop midline stereotypic hand movements, breathing irregularities, seizures, scoliosis, and disturbed sleeping patterns. Due to the broad variability of the phenotype in individuals with RTT, a proportion satisfy some, but not all of the necessary revised diagnostic criteria, and are classified as atypical RTT. Furthermore, individuals may be classified as RTT-like, exhibiting some clinical features associated with RTT, but not enough to be classified as either classic or atypical RTT (Ip, Mellios, & Sur, 2018). Approximately 97% of classic RTT and 86% of atypical RTT individuals have pathogenic variants in *MECP2* (Neul et al., 2014). A small proportion of individuals with clinical features overlapping with RTT have been reported with pathogenic variants in Cyclin dependent kinase like 5 (*CDKL5*, OMIM: 300203) and Forkhead box G1 (*FOXG1*, OMIM: 164874); however, a significant proportion of individuals clinically classified as RTT remain without a definite genetic diagnosis (Armani et al., 2012).

Functionally, MeCP2 binds to the methylated CpG dinucleotides and regulates epigenetic modifications via transcriptional repression and activation, chromatin remodelling and RNA splicing (Sanfeliu, Kaufmann, Gill, Guasoni, & Tropea, 2019). Such epigenetic modifications involving post-translational modification of chromosomal proteins, including histones, by reversible acetylation and/or methylation, control various facets of DNA transcription, replication and repair, processes which are critical for brain development and function (Radford, 2018). Histone deacetylases (HDACs) remove the acetyl group from histone tails, condensing the chromatin, and preventing transcription, whilst histone acetyltransferases (HATs) acetylate the conserved lysine residue of histones, resulting in chromatin unfolding and transcriptional activation (Portela & Esteller, 2010). Based on their homology, all 20 known HATs belong to one of the four families, including the MYST family (MOZ, Ybf2/Sas3, Sas2, and Tip60) whereby Lysine (K) acetyltransferase 6A (KAT6A/ MOZ/ MYST3) and its paralogue Lysine (K) acetyltransferase 6B (KAT6B/ MORF/ MYSR4) are involved in chromatin regulation by acetylating histone H3 at K-9 and K-14.

KAT6A/KAT6B forms a tetrameric complex with Runt-related transcription factor (RUNX), Bromodomain and PHD Finger Containing 1 (BRPF1), Esa1p-Associated Factor 6 (EAF6) and tumour suppressor ING5 proteins, and is important for differentiation of neuronal and hematopoietic stem cells (Pelletier, Champagne, Stifani, & Yang, 2002; Yang, 2015). Various animal and cellular model based studies have provided evidence that defects in epigenetic co-regulators KAT6A and KAT6B result in impaired neurogenesis and craniofacial abnormalities, which may be relevant to developmental disorders in humans (Katsumoto et al., 2006; Kong et al., 2014; Merson et al., 2006; Perez-Campo et al., 2014; Thomas, Voss, Chowdhury, & Gruss, 2000).

Protein truncating variants in *KAT6A* (OMIM: 6162680) and *KAT6B* (OMIM: 605880) have been reported in individuals with syndromic developmental delay and/or intellectual disability (Arboleda et al., 2015; Campeau et al., 2012; Tham et al., 2015). Heterozygous *de novo* variants in *KAT6B* have been found to be associated with Say-Barber-Biesecker-Young-Simpson syndrome (OMIM: 603736) characterized by intellectual disability as well as with congenital anomalies (Clayton-Smith et al., 2011) and genitopatellar gyndrome (OMIM: 606170), a rare condition with intellectual disability, genital anomalies and craniofacial defects (Campeau et al., 2012; Simpson et al., 2012). In addition, *de novo* heterozygous variants in *KAT6A* have previously been associated with *KAT6A*-related intellectual disability syndrome, a rare neurodevelopmental disorder with variable clinical features depending upon the location of the pathogenic variant. Although over 50 protein-truncating, six missense and four splicing variants feature across the whole *KAT6A* gene, the severity of clinical symptoms including intellectual disability, speech delay, microcephaly, cardiac anomalies, and gastrointestinal complications, has been observed to be more severe when the protein truncating pathogenic variants affect the last two exons of the gene (exon 16 and 17) (Kennedy et al., 2019).

Dysfunctional MeCP2, a transcriptional regulator, affects pathways resulting in NDDs. Moreover, there are several other gene products involved in transcriptional regulation pathway that when perturbed results in NDDs which have clinical features overlapping with RTT (Gold & Christodoulou, 2015; Kaur et al., 2019; Sanfeliu et al., 2019). Using next generation sequencing (NGS), we have identified five previously genetically undiagnosed RTT/ RTT-like individuals with pathogenic variants in *KAT6A*. Additionally, a thorough systematic reanalysis of 76 individuals with KAT6A syndrome already published (Kennedy et al., 2019) led to the re-classification of an additional two patients as RTT/RTT-like according to the Neul's revised diagnostic criteria (Neul et al.,

2010). Our findings highlight the importance of considering *KAT6A* as an alternative diagnosis in individuals who might appear to have RTT/RTT-like syndrome but are negative on *MECP2* testing.

2. Results

In this study we present seven unrelated individuals between the ages of 3 years 8 months and 21 years 6 months identified as RTT or RTT-like features, all of who have different *de novo* heterozygous, late truncating variants in *KAT6A*. Based on Neul's revised diagnostic criteria (Neul et al., 2010), we have classified individuals 1 and 5 as fitting a clinical diagnosis of atypical RTT. Five other individuals (2, 3, 4, 6 and 7) had an initial diagnosis of *KAT6A*-related intellectual disability but upon re-evaluation of their clinical features, they may be classified as RTT-like. Individual 6 and 7 have been reported previously by Kennedy et al 2019 ("Patient 13" and "Patient 43" respectively) (Kennedy et al., 2019). Clinical information used to classify the individuals is presented in Table 1. The seven individuals had several clinical features that are also encountered in KAT6A syndrome (Kennedy et al., 2019; Tham et al., 2015) which is detailed in Table S1.

2.1. Brief clinical summaries

Individual 1 was a 17 years and 5 month-old female with severe intellectual disability, development delay, microcephaly, growth retardation, behaviour problems (irritability), absent speech and hypsarrhythmia. She also exhibited regression, partial loss of acquired hand skills, gait abnormalities, stereotypic hand movements with extreme mouthing, severe hyperventilation, apnoea, bruxism, impaired sleep and spasticity. She also had small hands & feet with cool peripheries, inappropriate crying spills, and poor eye contact with intense eye communication along with scoliosis. According to Neul's criteria she has been classified as atypical RTT (3 main and all supportive criteria) (Table 1) (Neul et al., 2010). Genetic screening for genes associated with RTT (*MECP2*, *CDKL5* and *FOXG1*) was negative. Her key gastrointestinal features include feeding difficulties, gastroesophageal reflux and vomiting, severe constipation, and she also had intestinal malrotation with recurrent episodes of obstruction. She also had atrial and septal heart defects which have been surgically

corrected. She also had subtle dysmorphic features including wide sparse teeth, thick red lips, and high palate with normal eyes, ears and nose (Figure 2, Table S1).

Individual 2 was a 21 year and 6-month old Italian female with gross motor development delay, severe intellectual disability, autistic behaviour, poor social interactions, absent speech, abnormal muscle tone (hypotonia in infancy and peripheral hypertonia subsequently) and congenital microcephaly. She had feeding difficulties with delayed weaning during infancy. Other clinical features included mild hand stereotypies, gait abnormalities, kyphosis, inconsolable crying episodes, severe bruxism, and cold hands and feet. She did not exhibit any period of regression. She has been classified as RTT-like (2 main and 7 supportive using Neul's diagnostic criteria) (Table 1) but no pathogenic variants were identified on sequencing of *MECP2*, *CDKL5*, *UBE3A* and *FOXG1*. Her facial features included a broad bulbous nasal tip, bitemporal narrowing, prominent nasal bridge, short philtrum, and a thin upper lip and fleshy lower lip. She also had a narrow and high mouth palate, bilateral loss of the external third of her eyebrows, deep-set eyes, exotropia, high forehead, asymmetric face, prognathism and a prominent chin (Figure 2, Table S1). Along with recurrent urinary tract infections, she also had very frequent gastroesophageal reflux and severe constipation.

Individual 3 was a 3 years and 8 month old male with global development delay, absent speech, abnormal muscle tone, sagittal craniosynostosis, feeding difficulties, severe constipation, gastroesophageal reflux, bilateral undescended testes requiring orchidopexy, inguinal hernia repair, and chronic otitis media with effusion. He had cortical visual impairment that improved over time. He has been classified as RTT-like (1 main and 9 supportive criteria as per Neul's diagnostic criteria). He also exhibited dysmorphic features including a prominent nasal bridge with broad nasal tip, short philtrum and a thin upper lip, mild bitemporal narrowing, and widely spaced teeth. He had bilateral single palmar creases and small cold feet. (Figure 2 and Table S1).

Individual 4 was a 10 years and 4 month-old female with severe global development delay, intellectual disability and postnatal microcephaly. She was non-mobile, non-verbal and

displayed stereotypic hand movements including hand wringing/squeezing, clapping/tapping, mouthing and washing/rubbing. She also exhibited hyperventilation when excited, and had spasticity with hyperreflexia. She appeared to have high pain threshold and had alacrimia. She also exhibited tongue stereotypes and jerky movement of the limbs. Based upon her clinical features, we have classified her as RTT-like as she satisfied 2 main and 8 supportive criteria of Neul's diagnostic criteria (Table 1). She also exhibited facial dysmorphic features including thin upper lip, low set ears, prominent nasal bridge, smooth philtrum, high mouth palate and enamel hypoplasia (Figure 2, Table S1).

Individual 5 was a 12 years and 6 month-old male with global development delay and regression in association with infantile spasms. He was non-mobile, non-verbal and displayed stereotypic hand movements including hand wringing/squeezing, clapping/tapping, mouthing and washing/rubbing along, as well as inappropriate laughing and screaming spills. He was hypotonic, had an extremely high pain threshold and very intense eye communication (eye pointing). According to Neul's criteria he has been classified as atypical RTT (2 main and 9 supportive criteria) (Table 1). His dysmorphic features included thin upper lip, broad nasal bridge, high mouth palate, teeth abnormalities (Figure 2, Table S1), persistent foetal fat pads on his digits and prominent knuckles.

Individual 6 was a 5 years and 7 months old female with delayed developmental milestones particularly gross motor and speech development and intellectual disability. In addition, she had hypotonia, delayed social and emotional development, and an impaired sleep pattern with frequent nocturnal waking requiring melatonin treatment. We have classified her as RTT-like (1 main and 6 supportive criteria) (Table 1)

She also had feeding difficulties, and reflux as an infant. Constipation is a major ongoing issue requiring daily use of laxatives. She was also found to have mild left pulmonary artery stenosis, small patent foramen ovale (PFO) and patent ductus arteriosus (PDA). Other distinctive features included a left pre-auricular pit, delayed eruption of lower incisors that were widely spaced, unilateral supernumerary nipple and a raised median lingual ridge. Early episodes of recurrent impetigo and staphylococcal infections have largely resolved. Additional dysmorphic features have been previously reported (Figure 2, Table S1) (Kennedy et al., 2019).

Individual 7 was a 3 years and 11 month-old male with severe development delay, microcephaly, hypotonia, and absent speech. There was no regression noticed in his acquired skills. His stereotypic hand movements included constant tapping on chin, lip and forehead. Other features exhibited by him included bruxism while awake, impaired sleep patterns (waking several times in the night, and using a CPAP machine for sleep apnoea), minor scoliosis, small cold hands and feet, inappropriate laughing and screaming spells. He fulfilled 2 main and 5 supportive criteria of Neul's revised diagnostic criteria for RTT, thus we have classified him as RTT-like (Table 1).

He also had feeding difficulties since birth, gastroesophageal reflux and severe constipation. His facial dysmorphic features included narrow mouth, thin lips, low set ears, retrognathia and narrow forehead. In addition, he had cortical vision impairment, heart defects including autism spectrum disorder (ASD), PFO, PDA and leaky mitral valve

More detailed clinical summaries are available in the supplementary material.

2.2. Variant details

Using NGS and re-interrogation of reported individuals with *KAT6A*-related intellectual disability, seven different variants in *KAT6A* were identified in this cohort (Table 2). Four individuals had nonsense variants: NM_006766.5:c.3385C>T; NP_006757.2:p.(Arg1129*) in individual 1 (Family 1, II:3, Figure S1-A), NM_006766.5:c.3820G>T;NP_006757.2: p.(Glu1274*) in individual 2 (Family 2, II:1, Figure S1-B), NM_006766.5:c.3631_3632delGT;NP_006757.2:p.(Val1211*) in individual 5 (Family 5, II-2, Figure S1-E) and NM_006766.5: c.3661G>T;NP_006757.2:p.(Glu1221*) in individual 7 (Family 7, II:1 Figure S1-G). The other three individuals had frameshift variants that result in early truncation of protein translation: NM_006766.5:c.3399_3400dup;NP_006757.2:p.(Lys1134Argfs*14) in individual 3 (Family 3, II:2, Figure S1-C), NM_006766.5:c.3377delC;NP_006757.2:p.(Ser1126Phefs*8) in individual 4 (Family 4, individual II-2 Figure S1-D) and NM_006766.5:c.4254_4257del; NP_006757.2:p.(Glu1419Trpfs*12) in individual 6 (Family 6, II:1 Figure S1-F).

2.3. *In silico* analysis

All *KAT6A* variants were monoallelic and clustered in the last exon (exon 17) of *KAT6A* (Table 2). Of the seven different variants, four were novel (individuals 2 to 5) whilst the variant in individuals 1, 6 and 7 had been previously reported in individuals with significant developmental delay, cardiac defects and facial dysmorphism (Figure 1) (Arboleda et al., 2015; Kennedy et al., 2019). Segregation analysis confirmed variants in individuals 1, 2 and 3 to be *de novo* (Figure S1). Six of the variants (individuals 1 to 5 and case 7) were predicted to result in partial loss of the well-conserved glutamate/aspartate-rich acidic domain and complete loss of serine/methionine rich (Ser/Met) domain of KAT6A, whilst the variant in individual 6 resulted in complete loss of only the Ser/Met domain (Figure 1 and Figure S2). All seven *KAT6A* variants reported here are absent in the gnomAD population database. Moreover, pLI listed in gnomAD with score = 1.0, also highlighted that *KAT6A* is extremely intolerant towards loss of function variants. According to MutationTaster these variants are likely to be disease causing. Additionally, using the ACMG classification, all the variants are pathogenic (Table S2).

3. Discussion

Although variants in *KAT6A* have been previously associated with *KAT6A*-related intellectual disability spectrum of disorders, the association with RTT appears to have been previously unrecognized. We report seven individuals with pathogenic variants in *KAT6A* who have overlapping clinical features with RTT. Individual 1 and 2 were clinically diagnosed as atypical RTT and RTT-like respectively. Individuals 3 to 7 were initially diagnosed with *KAT6A*-related intellectual disability, of which individuals 6 and 7 have been previously published (Kennedy et al., 2019). We collected additional clinical information for these five individuals from their referring clinicians and re-classified these individuals as atypical RTT (individual 5) based upon Neul's diagnostic criteria (Neul et al., 2010) or RTT-like (individuals 3, 4, 6 and 7). All variants (4 novel and 3 previously reported) were predicted to be pathogenic and to cause C-terminal protein truncation, resulting in loss of the transactivation domain (glutamate/aspartate-rich acidic domain and Ser/Met domain) of the KAT6A protein. The nonsense mediated mRNA decay (NMD) pathway is an evolutionarily

conserved surveillance system that examines mRNA transcripts for premature termination codon errors and eliminates them from the transcriptome to avoid accumulation of deleterious proteins in the cell (Kurosaki, Popp, & Maquat, 2019). However, the efficiency of NMD largely depends upon the position of protein truncating variant within the gene. Based on the exon junction complex (EJC) model, if the premature termination codon is present at least 50 nucleotides upstream of the last exon-exon boundary, it might trigger NMD (Lindeboom, Supek, & Lehner, 2016). All the *KAT6A* variants discussed in this study are clustered in the last exon of the gene, and so are predicted to escape canonical NMD resulting in the generation of a truncated protein. Such a truncated protein may cause its effect through a dominant-negative or gain of function effect, potentially contributing to the phenotypic variability observed among individuals with *KAT6A* variants (Kennedy et al., 2019). This proposition is supported by a report that showed no significant decrease in the *KAT6A* mRNA levels in fibroblasts derived from two unrelated individuals with the p.(Arg1129*) variant compared to controls (Arboleda et al., 2015), indicating that NMD is not active in protein truncating variants affecting the last exon of *KAT6A*.

The variant in individual 1 [p.(Arg1129*)] is located at a CpG base and has been previously reported in eight affected individuals with MRD-32 (Arboleda et al., 2015), which is the most recurrent variant position in *KAT6A*. Consistent with eight previously published KAT6A syndrome individuals with p.(Arg1129*) variant, our individual 1 also exhibited intellectual disability, sleep disturbances, sleep apnoea, severe speech delay, feeding difficulties, gastroesophageal reflux, constipation, mild facial dysmorphism and dental abnormalities, and also had congenital cardiac defects (atrial septal defects and a ventricular septal defect). However, unlike reported individuals, she did not exhibit microcephaly, recurrent infections, or hypotonia and did not have a broad nasal tip. The proband also had infantile spasms with hypsarrhythmia, episodes of hyperventilation and breath holding which are not often reported in individuals with *KAT6A*-related intellectual disability, further highlighting the phenotypic variability among individuals with KAT6A variants. It was reported that this variant does not affect transcript stability (Arboleda et al., 2015), however analysis of global histone acetylation patterns found reduced H3K9 and enhanced H3K18 acetylation, suggesting dysregulated histone protein acetylation (Arboleda et al., 2015). Unfortunately, there was insufficient information in those publications to allow us to apply the Neul criteria with confidence.

All the reported individuals exhibited global development delay, severe intellectual disability, absent or severely delayed speech, abnormal muscle tone and sleep problems. In addition, all the patients had feeding difficulties, gastroesophageal reflux and constipation, which are common in individuals with late truncating pathogenic variants in *KAT6A* (Kennedy et al., 2019). Interestingly, individual 1 also had intestinal malrotation associated with recurrent bowel obstruction. All the individuals in this report showed discrete but overlapping dysmorphic facial features seen in individuals with *KAT6A*-related intellectual disability. While only half of the individuals showed broad nasal tip, bitemporal narrowing, high arched palate and philtrum defects, almost all of the individuals displayed thin upper lip, prominent nasal bridge and dental abnormalities. Four individuals (individual 1, 5, 6 and 7) exhibited cardiac defects, notably atrial septal defects, ventricular septal defects, PDA, PFO, leaky mitral valve or pulmonary artery stenosis. While individuals 1 and 4 did not showed any period of infections, all other five reported individuals (individuals 2, 3, 5, 6 and 7) showed recurrent episodes of different infections which is a common feature observed in *KAT6A*-related intellectual disability. Some individuals showed abnormalities in brain magnetic resonance imaging (MRI); delayed myelination (individuals 2 and 5), craniosynostosis (individuals 3 and 7) and Arnold-Chiari type I malformation (individual 5). Most of the reported individuals also had ocular abnormalities; myopia (individuals 1 and 4), cortical visual impairment (individuals 3, 5 and 7), or strabismus and/or had epicanthal folds (individuals 4, 5 and 6). Two males (individuals 3 and 5) also reported genitourinary problems including hydronephrosis and cryptorchidism.

Overall, although the reported RTT/RTT-like individuals had a number of overlapping *KAT6A*-associated features, they exhibited considerable variability in their phenotypic presentations. This may be partly explained by the histone acetylation function of KAT6A, which is regulated by additional epigenetic and environmental factors (Kennedy et al., 2019). Mice with a homozygous deletion of *Kat6a* die during embryogenesis or in the perinatal phase due to gastrointestinal abnormalities, cardiac defects, skeletal, haematological and immunological abnormalities (Katsumoto et al., 2006; Voss, Collin, Dixon, & Thomas, 2009; Voss et al., 2012). Of note, only approximately half of *Kat6a* heterozygous mice have developmental abnormalities of the palate, thymus, cardiovascular system and facial structures (Voss et al., 2012). These authors suggested that Kat6a may regulate the T-box transcription factor (*Tbx1*) locus. Further investigation of potential genetic and environmental factors suggested that abnormalities were due to either retinoic acid exposure or *Tbx1*

haploinsufficiency (Voss et al., 2012). It is possible that the broad phenotype of *KAT6A*-related intellectual disability and RTT individuals may be, in part due to pathogenic variants in *KAT6A*, but additional environmental or epigenetic factors may contribute to the complexity of the phenotype.

In conclusion, we have identified pathogenic protein truncating variants in *KAT6A* in seven individuals who could be classified as RTT/ RTT-like. We suggest that analysis of *KAT6A* should also be considered in the curation of genomic data of individuals with a clinical diagnosis of RTT, particularly when features overlapping with *KAT6A*-related intellectual disability are observed. Our work has highlighted the critical role of chromatin regulators in the pathogenesis of RTT, which could be considered as a potential for targeted drug therapies aiming at restoring a normal acetylation profile in RTT/ RTT-like individuals with *KAT6A* variants.

4. Material and Methods

4.1. Subjects and genetic analysis

All procedures used in this study were approved by the human research ethics committees of the participating institutes with written consents obtained from the legal guardians of each participant and in accordance with the Helsinki Declaration of 1975, as revised in 2000. The affected individuals with variants in *KAT6A* were recruited from Israel (individual 1), Italy (individual 2) and Australia (individual 3, individual 4 and individual 5) were assessed according to the revised diagnostic criteria for RTT (Neul et al., 2010). Genetic and clinical findings for individuals 6 and 7 were previously reported (Kennedy et al., 2019) and additional information regarding our RTT classification was obtained by directly contacting the referring clinician of the individual. Genomic DNA isolated from blood for all individuals was used for NGS as detailed in the Supplementary Material. The segregation of identified *KAT6A* variants was also confirmed by Sanger sequencing of affected individuals 1, 2 and 3 along with the parents as shown Supplementary Figure S1. Singleton exome sequencing was performed for individuals 4 and 5 in the VCGS NATA accredited laboratory, and the *de novo* status of the *KAT6A* variants identified was confirmed by parental Sanger sequencing, except for individuals 4 and 5, whose parents declined to have Sanger sequencing performed.

4.2. *In silico* predictions and variant classification

The mean allelic frequencies of the variants across all populations was determined using the Genome Aggregation Database (gnomAD) (Cambridge, MA; http://gnomad.broadinstitute.org/) (Karczewski et al., 2019). The pathogenicity of identified protein truncating variants was evaluated using the *in silico* tool MutationTaster (www.mutationtaster.org/) (Schwarz, Cooper, Schuelke, & Seelow, 2014) and pathogenicity was classified based on The American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015). In addition, the disease-specific allele frequency was determined from ClinVar, the public archive of human genetic variants (https://www.ncbi.nlm.nih.gov/clinvar/) (Landrum et al., 2018).

5. Acknowledgements

We thank all the family members for taking part in this study. The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program. The work conducted at University of Padova, Padova was supported by Italian Ministry of Health Young Investigator Grant GR-2011-02347754 to E.L.; Fondazione Istituto di Ricerca Pediatrica – Città della Speranza, Grant 18-04 to E.L. SK was the recipient of Research Training Program scholarship. We acknowledge the Victorian Clinical Genetics Services for performing WES. We also thank Chris Ieng and Damien Zammit for running the NGS data on Cpipe We would also like to acknowledge Vanessa Atkinson and Dr Karen Carpenter from Diagnostic Genomics Laboratory, PathWest for performing sequencing on individual 3. Recruitment and data collection for individual 7 was supported by the NHMRC Centre of Research Excellence in Speech and Language (#1116976).

6. References

Arboleda, V. A., Lee, H., Dorrani, N., Zadeh, N., Willis, M., Macmurdo, C. F., . . . Nelson, S. F. (2015). De novo nonsense mutations in KAT6A, a lysine acetyl-transferase gene, cause a syndrome including microcephaly and global developmental delay. *Am J Hum Genet, 96*(3), 498-506. doi:10.1016/j.ajhg.2015.01.017

Armani, R., Archer, H., Clarke, A., Vasudevan, P., Zweier, C., Ho, G., . . . Christodoulou, J. (2012). Transcription factor 4 and myocyte enhancer factor 2C mutations are not common causes of Rett syndrome. *Am J Med Genet A, 158A*(4), 713-719. doi:10.1002/ajmg.a.34206

Armstrong, A. H., Hangauer, J., Agazzi, H., Nunez, A., & Gieron-Korthals, M. (2010). Individuals with intellectual and developmental disabilities. In *Handbook of Pediatric Neuropsychology* (pp. 537-550). New York: Springer.

Campeau, P. M., Kim, J. C., Lu, J. T., Schwartzentruber, J. A., Abdul-Rahman, O. A., Schlaubitz, S., . . . Lee, B. H. (2012). Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet, 90*(2), 282-289. doi:10.1016/j.ajhg.2011.11.023

Clayton-Smith, J., O'Sullivan, J., Daly, S., Bhaskar, S., Day, R., Anderson, B., . . . Black, G. (2011). Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet, 89*(5), 675-681. doi:10.1016/j.ajhg.2011.10.008

Gold, W. A., & Christodoulou, J. (2015). The Utility of Next-Generation Sequencing in Gene Discovery for Mutation-Negative Patients with Rett Syndrome. *Front Cell Neurosci, 9*, 266. doi:10.3389/fncel.2015.00266

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

Ip, J. P. K., Mellios, N., & Sur, M. (2018). Rett syndrome: insights into genetic, molecular and circuit mechanisms. *Nat Rev Neurosci, 19*(6), 368-382. doi:10.1038/s41583-018-0006-3

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., . . . MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. 531210. doi:10.1101/531210 %J bioRxiv

Katsumoto, T., Aikawa, Y., Iwama, A., Ueda, S., Ichikawa, H., Ochiya, T., & Kitabayashi, I. (2006). MOZ is essential for maintenance of hematopoietic stem cells. *Genes Dev, 20*(10), 1321-1330. doi:10.1101/gad.1393106

Kaur, S., Van Bergen, N. J., Gold, W. A., Eggers, S., Lunke, S., White, S. M., . . . Christodoulou, J. (2019). Whole exome sequencing reveals a de novo missense variant in EEF1A2 in a Rett syndrome-like patient. *Clin Case Rep, 7*(12), 2476-2482. doi:10.1002/ccr3.2511

Kennedy, J., Goudie, D., Blair, E., Chandler, K., Joss, S., McKay, V., . . . Newbury-Ecob, R. (2019). KAT6A Syndrome: genotype-phenotype correlation in 76 patients with pathogenic KAT6A variants. *Genet Med, 21*(4), 850-860. doi:10.1038/s41436-018-0259-2

Kong, Y., Grimaldi, M., Curtin, E., Dougherty, M., Kaufman, C., White, R. M., . . . Liao, E. C. (2014). Neural crest development and craniofacial morphogenesis is coordinated by nitric oxide and histone acetylation. *Chem Biol, 21*(4), 488-501. doi:10.1016/j.chembiol.2014.02.013

Kurosaki, T., Popp, M. W., & Maquat, L. E. (2019). Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol, 20*(7), 406-420. doi:10.1038/s41580-019-0126-2

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., . . . Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res, 46*(D1), D1062-D1067. doi:10.1093/nar/gkx1153

Lindeboom, R. G., Supek, F., & Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat Genet, 48*(10), 1112-1118. doi:10.1038/ng.3664

Merson, T. D., Dixon, M. P., Collin, C., Rietze, R. L., Bartlett, P. F., Thomas, T., & Voss, A. K. (2006). The transcriptional coactivator Querkopf controls adult neurogenesis. *J Neurosci, 26*(44), 11359-11370. doi:10.1523/JNEUROSCI.2247-06.2006

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

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

Pelletier, N., Champagne, N., Stifani, S., & Yang, X. J. (2002). MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. *Oncogene, 21*(17), 2729-2740. doi:10.1038/sj.onc.1205367

Perez-Campo, F. M., Costa, G., Lie, A. L. M., Stifani, S., Kouskoff, V., & Lacaud, G. (2014). MOZ-mediated repression of p16(INK) (4) (a) is critical for the self-renewal of neural and hematopoietic stem cells. *Stem Cells, 32*(6), 1591-1601. doi:10.1002/stem.1606

Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nat Biotechnol, 28*(10), 1057-1068. doi:10.1038/nbt.1685

Radford, E. J. (2018). An Introduction to Epigenetic Mechanisms. *Prog Mol Biol Transl Sci, 158*, 29-48. doi:10.1016/bs.pmbts.2018.04.002

Rett, A. (1966). [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien Med Wochenschr, 116*(37), 723-726.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Sanfeliu, A., Kaufmann, W. E., Gill, M., Guasoni, P., & Tropea, D. (2019). Transcriptomic Studies in Mouse Models of Rett Syndrome: A Review. *Neuroscience, 413*, 183-205. doi:10.1016/j.neuroscience.2019.06.013

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods, 11*(4), 361-362. doi:10.1038/nmeth.2890

Simpson, M. A., Deshpande, C., Dafou, D., Vissers, L. E., Woollard, W. J., Holder, S. E., . . . Trembath, R. C. (2012). De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. *Am J Hum Genet, 90*(2), 290-294. doi:10.1016/j.ajhg.2011.11.024

Tham, E., Lindstrand, A., Santani, A., Malmgren, H., Nesbitt, A., Dubbs, H. A., . . . Nordgren, A. (2015). Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. *Am J Hum Genet, 96*(3), 507-513. doi:10.1016/j.ajhg.2015.01.016

Thomas, T., Voss, A. K., Chowdhury, K., & Gruss, P. (2000). Querkopf, a MYST family histone acetyltransferase, is required for normal cerebral cortex development. *Development, 127*(12), 2537-2548.

Voss, A. K., Collin, C., Dixon, M. P., & Thomas, T. (2009). Moz and retinoic acid coordinately regulate H3K9 acetylation, Hox gene expression, and segment identity. *Dev Cell, 17*(5), 674-686. doi:10.1016/j.devcel.2009.10.006

Voss, A. K., Vanyai, H. K., Collin, C., Dixon, M. P., McLennan, T. J., Sheikh, B. N., . . . Thomas, T. (2012). MOZ regulates the Tbx1 locus, and Moz mutation partially phenocopies DiGeorge syndrome. *Dev Cell, 23*(3), 652-663. doi:10.1016/j.devcel.2012.07.010

Yang, X. J. (2015). MOZ and MORF acetyltransferases: Molecular interaction, animal development and human disease. *Biochim Biophys Acta, 1853*(8), 1818-1826. doi:10.1016/j.bbamcr.2015.04.014

7. Tables

Category	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
Age (as of Jan 2020)	17 years and 5 month	21 years and 6 months	3 years and 8 months	10 years and 4 months	12 years and 6 months	6 years and 7 months	3 years and 11 months
Gender	Female	Female	Male	Female	Male	Female	Male
Ancestry	Ashkenazi Jewish	Italian	Caucasian/ Portuguese	European	European	Anglo-Celtic	European
Main criteria							
Partial/ complete loss of acquired purposeful hand skills	+	-	-	-	-	-	-
Partial/ complete loss of acquired spoken language	-	-	-	-	-	-	-
Gait abnormalities: Impaired/absence of ability	+	+	+	+	+	+	+
Stereotypic hand movements [a]	+	+	-	+	+	-[&]	+
Exclusion criteria							
Brain injury [b]	-	-	-	-	-	-	-
Grossly abnormal psychomotor development (first 6 months)	+	+	+	+	+	+	+
Supportive criteria for atypical RTT							
Breathing disturbances when awake	+	-	+	+	+	-	-
Bruxism when awake	+	+	+	+	+	-	+
Impaired sleep pattern	+	+	+	+	+	+	+
Abnormal muscle tone	+	+	+	+	+	+	+
Peripheral vasomotor disturbances	+	+	+	-	-	-	-
Scoliosis/kyphosis	+	+	+	+	+	-	+
Growth retardation	+	-	-	+	-	-	-
Small cold hands and feet	+	+	+	+	-	+	+

Inappropriate laughing/screaming spells	+	-	+	-	+	+	+
Diminished response to pain	+	+	+	+	+	+	-
Intense eye communication – "eye pointing"	+	-	-	-	+	-	-
Required for classic (typical) RTT							
A period of regression followed by recovery or stabilization	+	-	-	-	+	+[#]	-
All main criteria and all exclusion criteria	-	-	-	-	-	-	-
Required for atypical or variant RTT							
A period of regression followed by recovery or stabilization	+	-	-	-	+	+	-
At least 2 of the 4 main criteria	+ (3/4)	+ (2/4)	- (1/4)	+ (2/4)	+ (2/4)	+ (1/4)	+ (2/4)
5 out of 11 supportive criteria	+ (11/11)	+ (7/11)	+ (9/11)	+ (8/11)	+ (8/11)	+ (5/11)	+ (6/11)
Classification	**Atypical RTT**	**RTT-like**	**RTT-like**	**RTT-like**	**Atypical RTT**	**RTT-like**	**RTT-like**

Table 1: Assessment of *KAT6A* individuals using the revised diagnostic criteria (Neul et al 2010).

a: hand wringing/squeezing, clapping/ tapping, mouthing and washing/rubbing automatisms; b: secondary to trauma (peri- or postnatally), neurometabolic disease, or severe infection that causes neurological problem; #: Individual 6 rolled from front to back at one month of age and then stopped at 3 months of age; &: mouthing and recurrent arm flapping

Category	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
Genetic screening method	Singleton; WES	Trio; WES	Singleton; Targeted sequencing	Singleton; WES	Singleton; WES	Singleton; WES	Singleton; WES
Variant	chr8:41792353G>A c.3385C>T; p.(Arg1129*)	chr8:41791918C>A c.3820G>T p.(Glu1274*)	chr8:41792338_41792339dup c.3399_3400dup p.(Lys1134Argfs*14)	chr8:41792361delC c.3377delC p.(Ser1126Phefs*8)	chr8:41792106_41792107delGT c.3631_3632delGT p.(Val1211*)	chr8:41791481_41791484del c.4254_4257del p.(Glu1419Trpfs*12)	chr8:41792077C>A c.3661G>T p.(Glu1221*)
Zygosity	het	het	het	het	het	het	het
Segregation	*de novo*	*de novo*	*de novo*	*Not investigated*^	*Not investigated*^	*de novo*$	*de novo*$
loss of conserved domains	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Novel or reported	Reported Arboleda et al (2015)	Novel	Novel	Novel	Novel	Reported Kennedy et al (2018); (P 13)	Reported Kennedy et al (2018); (P 43)
Affected exon	exon (17/17)	exon (17/17)	exon (17/17)	exon (17/17)	exon (17/17)	exon (17/17)	exon (17/17)
Location of variant	acidic domain and Ser/Met domain	acidic domain and Ser/Met domain	acidic domain and Ser/Met domain	acidic domain and Ser/Met domain	acidic domain and Ser/Met domain	Ser/Met domain only	acidic domain and Ser/Met domain
***in silico* (Mutation Taster)**	disease causing	disease causing	disease causing	disease causing	disease causing	disease causing	disease causing
gnomAD	Absent	Absent	Absent	Absent	Absent	Absent	Absent
ClinVar	Present	Absent	Absent	Absent	Absent	Present	Present
ACMG classification	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic

Table 2: *KAT6A* variant analysis for RTT/RTT-like individuals

GenBank transcript ID: NM_006766.5 (longest isoform), position: hg19:chr8:41,786,997-41,909,505, size: 122,509 bp, total exon count: 17; NCBI protein ID: NP_006757.2 (longest isoform), protein domains: N-term region of ENOK, MOZ or MORF (NEMM domain; aa 1–206), plant homeodomain (PHD domain; 207–313), Histone acetyltransferases (HAT domain; aa 314–787), glutamate/aspartate-rich acidic domain (788–1414), and serine- and methionine-rich transactivation domain (Ser/Met domain; aa 1414–2004). ^: testing was done in NATA accredited lab. The Integrated Genomic Viewer (IGV) reads showed high quality variants; $: variants already reported to be *de novo* (Kennedy et al., 2019)

8. Figure legends

Figure 1: Graphic representation showing KAT6A domain structure and location of identified KAT6A variants.

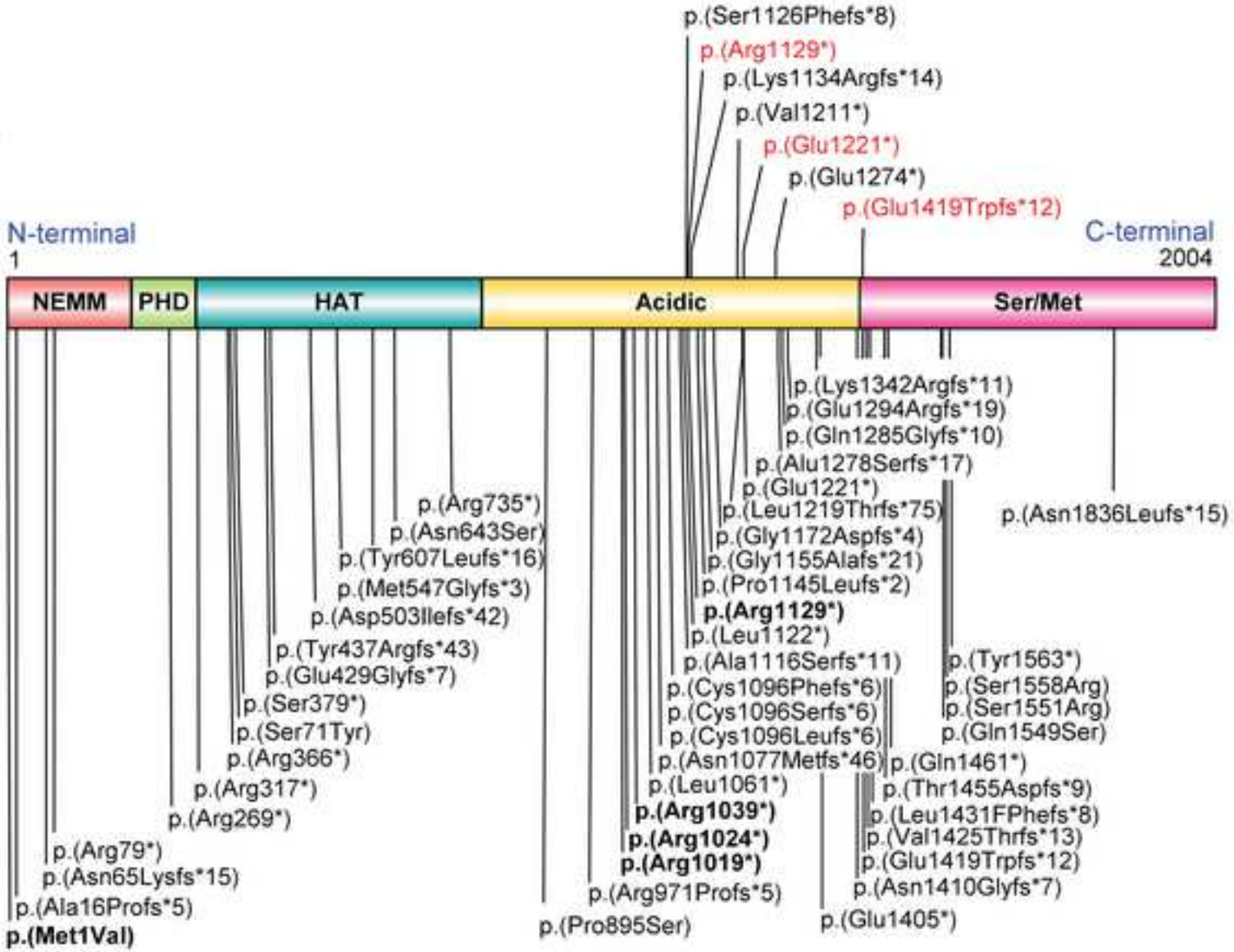
Schematic showing structure of key domains in KAT6A using protein ID NP_006757.2 (1-2004aa). N-terminal part of Enok, MOZ or MORF (NEMM) domain (1-206 aa; red; DNA binding with preference to CpG motifs), PHD (plant homeodomain) domain (207-313 aa; light green; histone binding), histone acetyl-transferase (HAT) domain (314-787 aa; dark green; acetyltransferase), a glutamate/aspartate-rich acidic domain (788-1414 aa, yellow), and a serine- and methionine-rich transactivation domain (Ser/Met) domain (1414-2004 aa, pink; transcriptional activation). The variants identified in the individuals with RTT/ RTT-like clinical phenotype are displayed above the KAT6A protein domain structure, and previously reported variants identified in individuals with *KAT6A*-related intellectual disability (Kennedy et al., 2019) are shown below the protein domain structure. Variants in red represent previously reported KAT6A variants observed in the RTT/ RTT-like patients in this study. Five variants in bold font have been found in multiple individuals with *KAT6A*-related intellectual disability: p.(Met1Val)(n=2), p.(Arg1019*)(n=2), p.(Arg1024*)(n=2), p.(Arg1039*)(n=2) and p.(Arg1129*)(n=8). aa = amino acid.

Figure 2: Clinical photographs of six individuals with RTT/RTT-like features and variants in *KAT6A*.

RTT/ RTT-like individuals with *KAT6A* variants show distinct facial features. **Individual 2** at approximately 9 months (A), approximately 5 years (B), at 13 years and 5 months (C) and at 17 years and 3 months (D). **Individual 3** at 6 months and 24 days (E, F), at 1 year old (G, H) and 3 years and 8 months (I). **Individual 4** at 8 years (J). **Individual 5** at 2 days (K), 4 years (L), 8 years and 6 months (M), 10 years (N), 11 years (O), **Individual 6** at 4 months (P), 6 years and 2 months (Q) and **Individual 7** at 1 year (R), 2 years (S), 3 years (T) and 3 years and 11 months (U). Individual 6 and Individual 7 have been previously published in (Kennedy et al., 2019). Clinical photos of these individuals show mild dysmorphic facial features as observed in individuals with *KAT6A*-related intellectual disability. The common facial features among all the affected individuals include thin upper lip, fleshy lower lip and prominent nasal bridge, however other features such as broad nasal tip, bitemporal narrowing and philtrum defects are only observed in the subset of individuals.

Variants identified in RTT/RTT-like patients

Previously reported variants identified in KAT6A patients




A Individual 2: ~ 9 months
B Individual 2: ~ 5 years
C Individual 2: 13 years 5 months
D Individual 2: 17 years 3 months
E Individual 3: 6 months 24 days
F Individual 3: 6 months 24 days
G Individual 3: 1 year
H Individual 3: 1 year
I Individual 3: 3 years 8 months
J Individual 4: 8 years
K Individual 5: 2 days
L Individual 5: 4 years
M Individual 5: 8 years 6 months
N Individual 5: 10 year
O Individual 5: 11 years
P Individual 6: 4 months
Q Individual 6: 6 years 2 months
R Individual 7: 1 year
S Individual 7: 2 years
T Individual 7: 3 years
U Individual 7: 3 years 11 months

1. Supplementary material

Expanding the genetic landscape of Rett syndrome to include Lysine (K) acetyltransferase 6A (*KAT6A*)

Authors

Simranpreet Kaur [a,b], Nicole J. Van Bergen [a,b], Bruria Ben-Zeev [c,d,e], Emanuela Leonardi [f,g], Tiong Y. Tan [b,h], David Coman [i,j,k], Benjamin Kamien [l,m], Susan M. White [b,h], Miya St John [n,o], Dean Phelan [h], Kristin Rigbye [h], Sze Chern Lim [h], Michelle C. Torres [h], Melanie Marty [h], Elena Savva [h], Teresa Zhao [h], Sean Massey [a], Alessandra Murgia [f,g], Wendy A. Gold [p,q,r], John Christodoulou [a,b,h,p]

1.1. Detailed clinical summaries

Individual 1 was a 17 years and 5 months old Ashkenazi Jewish female who can speak only one word and preferred objects close to face. She exhibited severe delay in gross motor skills. She had a history of seizures first triggered after hypsarrhythmia which are not currently ongoing. She could walk few steps and motor regression was absent. Her skin colour was noted to be very fair which was distinct from the family and used to get turned to bluish colour very easily.

Previous testing of *MECP2*, *CDKL5* and *FOXG1* were normal. In 2010, WES performed in Italy found a variant in *SLC2A1* (solute carrier family 2) that codes for the major glucose transporter in brain. However, the CSF glucose was normal. Thus, the proband was considered to be a mutation negative RTT individual.

Individual 2 was a 21 year old Italian female firstborn to non-consanguineous Italian parents. The mother was 32 and the father was 36 years old. The child was born at 36 weeks of gestation by caesarean section due to podalic presentation; birth weight 2.470 kg, length 46.5 cm and CC 32.5 cm. She had a regular perinatal period. Social smile was observed at three

months of life, however the baby was hypotonic and did not reach head control until 6-7 months. She could sit at 1 year of age, crawled at 2, and walked independently at 4 years of age. In infancy she was frequently agitated with inconsolable crying, not interested in the environment; interaction with her parents was minimal. She always suffered from severe constipation, had feeding difficulties and an altered sleeping pattern. She never started babbling and still she can only produce guttural sounds.

At 4 months of life she came to medical attention for global developmental delay and suspected hypo-vision. Cerebral ultrasound, EEG, and ECG resulted normal; brain MRI revealed a myelination delay in the temporal lobes and semioval center. Visual evoked potentials performed at 6 months revealed a pathological retino-cortical transmission and altered cortical pattern in response to visual stimuli.

From the age of 4, the girl started to manifest nocturnal episodes of noisy swallowing with intense salivation, lasting a few minutes, sometimes associated with vomiting. However, the EEG periodically performed in wakefulness and during sleep presented a normal pattern. In the contest of a severe neurodevelopmental condition, several dysmorphic features, noticed since infancy, suggested a possible diagnosis of Angelman syndrome: microcephaly, exotropia, epicanthic folds, wide mouth, together with behavioural traits (such as the fascination for water bottles and reflecting surfaces), absent speech and ataxic gate. The child was also diagnosed with farsightedness and astigmatism.

At clinical evaluation, at the age of 12 years, she presented severe intellectual deficit and continued to have poor social interaction with clear ASD traits: she kept presenting stereotypies and repetitive motor patterns (fascination with water bottles; moving a water bottle from one hand to another). However, she was referred as having motivated smile and seemed to be less irritable. She could somehow eat independently but was not toilet trained. Dysautonomic alterations, together with severe constipation included cold extremities. Since infancy she kept suffering from frequent urinary infections. Despite an improved sleeping pattern, nocturnal episodes of swallowing persisted. Speech was still absent, and she showed ataxic gate with hyperflexed trunk. On physical examination, she presented microcephaly (OFC: 50cm, 3°p) and kyphosis. Facial dysmorphic features became coarser with age: asymmetric face, high forefront, eyebrows lacking the external third, deep set eyes, hypertelorism, high nasal bridge with a prominent nose characterized by bulbous tip, short

philtrum. wide mouth with thin upper lip, prominent chin with prognathism. In addition, the proband also exhibited drug-induced oculogyric crises.

Previous genetic tests

First diagnostic examinations of individual 2, which included conventional karyotyping and sub-telomeric FISH analysis, comparative genomic hybridization (CGH) (75kb resolution), Fragile X testing and methylation analysis of the Prader-Willi/Angelman critical region, all resulted negative. Sanger sequencing and real time PCR of the corresponding exonic regions were performed to exclude alterations of the *MECP2*, *UBE3A*, and *CDKL5* genes. Target sequencing of a 74 gene panel (including *FOXG1*) for Intellectual Disability and Autism Spectrum Disorder was also resulted negative (Aspromonte et al., 2019).

Individual 3 was a 3 years and 8 months old Caucasian/ Portuguese male who was a second child to non-consanguineous healthy parents who had subsequently given birth to healthy twins. After birth, the proband did not cry for the first 10 days, but later he was constantly crying with a high-pitched cry and was an extremely unsettled baby. He did not give a social smile until age 5 months. He had bilateral undescended testes with right inguinal hernia repair. Fisted hands were noticed until age of 7 months but overall moves hand well and about to develop pincer grip. Sagittal craniosynostosis was noted and a spring cranioplasty performed at age three and a half months. He was noticed to have no visual engagement and review by ophthalmology found no structural eye abnormalities. A diagnosis of cortical visual inattention was made. Griffith's testing at age 13 months suggested a developmental age of approximately 4 months. Examination at age 13 months showed weight 10.3kg (50th percentile), length 78cm (~75th percentile), and head circumference of 45.8cm (10th percentile). Mild bitemporal narrowing, full nasal tip, bright blue eyes, and bilateral simian creases along with tapered fingers were also noted. Investigations included CT skull at age 1 year that demonstrated isolated sagittal craniosynostosis.

His other clinical features included intellectual disability, absent speech, gait abnormalities (at 3 year and 2 months - could crawl, pull to stand and walk up to two steps before falling down), mild scoliosis, hypotonia (trunk), breathing irregularities (strange rapid and then normal pattern of breathing with frequent breath holding episodes), bruxism (at 2 years of

age), abnormal muscle tone (dystonia in legs and mild hypertonia in trunk and upper body) and inappropriate laughing/ screaming spells. From birth until 1 year of age, he had difficulties falling asleep. Even though, he was given melatonin to initiate sleep, an extra dose of gabapentin and/or chloral hydrate was also required to settle irritability at night. Clonidine was found to be ineffective in his individual. He often slept for 8-10 hours with intermittent waking up. In the past, he had five episodes of generalised clonic tonic seizures and possibly absence and gelastic seizures as well (currently being investigated). He may had choreiform movements and dyskinesia when unwell and bilateral single palmar creases.

He also exhibited diminished response to pain and peripheral vasomotor disturbances (cold and blue extremities) and glue ear, tongue stereotypes noticed in this individual included constant biting, chewing, flipping and/or bending. As per the revised Neul's clinical diagnostic criteria he was classified as RTT-like (1 main and 9 supportive criteria).

Individual 3 showed oral aversions and did not tolerate certain food textures (only tolerated wet food). He had frequent and significant gastroesophageal reflux (settled with omeprazole) and severe constipation (required enemas weekly and Movicol daily). He suffered from frequent infections including lung, eye, ear and urinary tract infections along with mild renal reflux. Circumcision was performed to improve the frequencies of infections.

Other normal investigations

Chromosomal microarray (Illumina HumanCytoSNP-12 BeadChip (approximately 300,000 markers) and Karyostudio software version 1.4 [genome assembly GRCh37/hg19]) was reported as normal at age 9 months. Brain MRI and hip x-rays at age 11 months were unremarkable. Urine metabolic screen at age 15 months was unremarkable. Creatine kinase, ammonia, pyruvate and lactate, liver function tests, plasma amino acids, plasma acylcarnitine profile, lysosomal enzyme analysis, cerebrospinal fluid amino acid analysis and neurotransmitter analysis were unremarkable. Echocardiogram at age 1 year and 6 months was unremarkable. Rectal biopsy at age 1 year and 7 months was unremarkable. Micturating cystourethrogram at age 1 year and 11 months was reported as normal. Barium swallow at age 2 year and 1 month was reported as normal with no evidence of malrotation. The dimercaptosuccinic acid renal scan was normal at age 2 year and 2 month. Repeat brain MRI at the age of 3 year and 3 months was unremarkable aside from pansinusitis and bilateral mastoid air cell effusions.

Individual 4 was a 10 years and 4 months old European female who was the product of a natural conception. There were concerns raised with the triple test in early pregnancy and her mother went on to have an amniocentesis which showed normal chromosomes. There were normal fetal movements and the fetal morphology screen was normal. She was born via elective repeat caesarean section at 38 weeks gestation. There were no concerns noted at birth. She did have difficulties with breast feeding but bottle fed well. She had significant problems with tongue thrusting and bulbar dysfunction necessitating nasogastric tube (NGT) feeds from 6 weeks of age. Concerns about her development were first noted around 6 months of age. She had back arching. She was also noted to have torticollis. She had fisting and was noted to be hypertonic. She also had bilateral esotropia and was reviewed by the ophthalmologist. When first reviewed by a neurologist at 11 months of age she could roll from prone to supine but not back the other way. She had poor head control and could not sit independently. Regarding fine motor control, she had recently started opening her fists and begun grabbing some objects. She had some recognition of family members. She had no speech. Her tone had improved since she started treatment with Baclofen but she remained hypertonic. She had abnormal eye movements with intermittent esotropia of both eyes. She would fix and follow at times but this was not consistent. Her pupils were equal and reactive to light. She had bulbar dysfunction with notable tongue thrusting and a poor suck. She had evidence of grasp reflex but would not hold. She had generalized hypertonia with her head in the same plane as her body when pulling to sit. She was hyperreflexic with cross adductors present. There was no ankle clonus and her plantar responses were extensor bilaterally. At 3 years of age she would eat finger food. At this time she started to show affection and have more signs of social-emotional development, and increasing her vocalizations but had no formal language. At 10 years of age she was non-verbal and unable to walk. She used her hands and arms to handle objects but without purpose. Her other features included mild and persistent increase in creatine kinase, anhydrosis, difficulty in sleep initiation and maintenance. Other normal genetic and metabolic investigation included chromosomal microarray, MRI, transferrin isoforms and Apolipoprotein C3, Congenital Disorders of Glycosylation gene panel (Katholieke Universiteit Leuven), white cell enzyme and mitochondrial DNA analysis. The mitochondrial respiratory chain enzyme analysis showed normal to elevated levels when expressed relative to protein, relative to citrate synthase and

relative to complex II which are not suggestive of a respiratory chain defect. In absence of seizures, *MECP2* testing was not performed.

Individual 5 was a 12 years and 6 months old European male with significant development delay was born to non-consanguineous healthy parents. Pregnancy was full term with an elective caesarean section. Birth measurements for the child were within the average to above average range: length 54 cm (95th percentile), weight 3505 grams (50th percentile) and head circumference 35.5 cm (50th percentile). Infantile spasms were noted as early as 6 months of age with the child experiencing 100-200 seizures a day at a frequency of 1-2 seizures every 15 minutes. At 5 years of age his parents reported a well-documented history of marked mood lability associated with "manic-like symptoms", followed by what appeared to be irritability or low mood. The parents documented these to occur in cyclic episodes over a 3-4-week period. "Manic" episodes were characterized by marked excitability, poor sleep, episodes of giggling and hyperactivity. "Depressive" episodes were characterized generally as a withdrawn state with increased sleep, reduced responsiveness to external stimuli, lack of interest in his environment and decreased vocalizations. "Manic" and "depressive" periods generally lasted for one week each. No precipitants are identified and the episodes do not appear to be related to any seizure-like or prodromal symptoms. "Rapid cycling" between the two mood states crested a significant burden for the family. There was no family history of early onset or adult onset bipolar affective disorder or psychotic illnesses. EEGs have documented no seizure episodes associated with these episodes. A significant reduction in the frequency and intensity of these mood swings was achieved with the introduction of the mood stabilizing agent, sodium valproate. By 12 years of age he was nonverbal, with intense stare. His hands demonstrated stereotypical mouthing movements of objects, but he was unable to use them for activities of daily living, for which he was fully dependant on his parents and carers. His multisystem phenotype is characterised by Arnold-Chiari type I malformation, cortical visual impairment, cryptorchidism, persistent foetal fat pads on his digits, prominent knuckles, dysmorphic facial features, brittle hair in the form of trichorrhexis nodosa and trichoptilosis, atrial septal defect, "bipolar-like" mood disturbances controlled with mood stabilisers and recurrent cyclic mouth ulcers that responded to oral steroids. *MECP2* testing was not performed. The CGH array was normal. The SNP array detected 1q21.1 duplication (NCBIv36/hg18, 143,863,632-144,544,342) which was also found in several other clinically normal family members from three generations. As per the discussion

with the 1q21 consortium, this duplication would not be the sole pathology owing to the degree of clinical variance between the proband and his mother.

Individual 6 was a 6 years and 7 months old Anglo-Celtic female who was born to unrelated parents. The proband was extremely inactive in the womb and would usually only respond to light. She was able to roll from front to back at one month of age but then stopped at 3 months, with no further attainment of skills until 7 months of age. Regression in speech was also noticed when she stopped verbalizing after learning to babble at 11.5 months of age followed by recovery. The proband would repetitively say a word often for a short period before completely stopping that particular word. She had started taking steps on a trampoline at 2 year and 5 months of age. Her legs were always stiff, and toes pointed. She also exhibited several hand stereotypic movements such as mouthing, arm flapping across body and sometimes up and down from elbow particularly while in bed. Her body temperature was also not regulated properly.

Individual 7 was a 3 years and 11 months old European male born to unrelated parents through pregnancy complicated by gestational diabetes and low amniotic fluid. He was born at 38.5 weeks of gestation with birth weight of 2.98 kg. There were no complications during the birth but signs of anoxia, hypertonia, inability to nurse and laryngomalacia were observed within the first 4 weeks of life. His laryngomalacia involved a high-pitched wheezing sound when drinking from a bottle. This was corrected using a feeding tube and supraglottoplasy surgery at 7 months of age. The individual had normal hearing without any history of ear infections. He could sit unsupported at 13 months of age but still could not crawl or walk.

Individual 7 had frequent episodes of infections and his skin problem included eczema. He also had severe constipation managed by Miralax. His regular medication included carbamazepine for a mood dysregulation disorder and carnitine for possible mitochondrial deficiency along with other supplements such as vitamin B, C, D, E CoQ10, magnesium, calcium and fish oil. Individual 7 underwent several surgeries throughout his first year of life including cranial vault remodelling surgery (craniosynostosis) at 4 months for sagittal synostosis, supraglottoplasty at age 6 months for laryngomalacia, double orchidopexy for undescended testes and a hernia repair at 7 months, and the placement of a NGT at 8 months. The results of an EEG and MRI at 7 months of age were unremarkable.

1.2. NGS details

Individual 1

In this individual, singleton WES was performed on genomic DNA extracted from blood using previously described method (Kaur et al., 2019). Briefly, libraries were prepared and sequenced at the Translational Genomics Unit, Victorian Clinical Genetic Services (VCGS) using Illumina's NextSeq 500 sequencer. Raw sequencing data was converted to FASTQ format using Illumina's bcl2fastq2 converter (v2.17.1.14) and further processed using Cpipe (http://cpipeline.org/) (Sadedin et al., 2015), in order to generate annotated variant calls within the target region (coding exons +/- 2bp), via alignment to the reference genome (GRCh37). All variants found via Cpipe (n=25,534) were uploaded onto Leiden Open (source) Variation Database (LOVD) prior to further analysis (Fokkema et al., 2011). First, the RTT associated genes were screened to examine any previously missed variants (Christodoulou, Grimm, Maher, & Bennetts, 2003; Krishnaraj, Ho, & Christodoulou, 2017). Variant filtering was performed against the VCGS pre-curated intellectual disability (ID) gene list (https://www.vcgs.org.au/sites/default/files/media/TGW024_genelist_V3_0.pdf; n=1064). The remaining 615 variants were further prioritised based upon the impact of the variant in the order of early stop gain, frameshift and missense variants.

The sequence encompassing the identified KAT6A variant was amplified in proband and both the parents using DNA isolated from blood. The forward and reverse primers designed by Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) were 5'- tgtaagaccttgtaacctgtctttg-3'and 5'-attggtttgcggctcttg-3' respectively.

Individual 2

Written informed consent for WES was obtained from the individual's parents by clinical provider. This study was approved by the Local Ethics Committee, University-Hospital of Padova, Italy. Genomic DNA was extracted from leukocytes after blood cell isolation from proband and both parents using Wizard genomic DNA Promega Kit (Promega Corporation). Isolated DNA was sent to Biodiversa s.r.l., Trento, Italy. Exonic, UTR, and regulatory regions of the genome were isolated by capture using the Agilent SureSelect Human All

Exon V6 + UTR (71Mb). DNA was fragmented and ligated to the Illumina multiplexing paired-end adapters, followed by PCR amplification using primers with sequencing barcodes. Amplified products were then captured by biotinylated RNA library baits in solution following the manufacturer's instructions. Bound DNA was isolated with streptavidin-coated beads and reamplified. Paired-end reads 150 bp were generated on an Illumina HiSeq 4000 platform (Illumina, San Diego, CA). The output data were mapped to the reference human genome sequence GRCh37/hg19 using the BWA package. Quality control of sequencing data was obtained with Picard: >95% of reads aligned to target with Q20 base call quality, >85% target base covered at >30X, mean coverage of target bases >100. The Genome Analysis Toolkit (GATK) was used for base quality score recalibration and SNP and short INDEL calling.

Variants were annotated with Annovar (Wang, Li, & Hakonarson, 2010) and *de novo,* autosomal recessive, and X-linked variants were identified. Common population variants were excluded based on minor allele frequency (>0.5%) evaluated in GnomAD database, 1000 Genome Project (Genomes Project et al., 2015), and NHLBI exome sequencing Project (ESP, https://evs.gs.washington.edu/EVS/). An in-house pipeline was created to build a local database to exclude variants frequently identified in our cohort. Variants were classified by Intervar (Li & Wang, 2017) based on the system for grading evidence for pathogenicity by the ACMG guidelines. Pathogenicity predictions were assessed with the consensus among the computational tools provided by Annovar and by *in silico* analysis of the possible functional and structural protein effects of the candidate variants. Particular attention was made for variants in genes already known to be associated with phenotypes sharing common aspects with RTT or in genes involved in pathways relevant for neurodevelopment. Candidate sequence variants were confirmed with Sanger sequencing using primer pair 5'-gaggcagaagaggctgaaga-3' (forward) and 5'-aatccagctcggtttcctct-3' (reverse).

Individual 3

NGS was performed on genomic DNA extracted from blood using Illumina TruSight One sequencing kit (FC-141-1006) on Illumina Mi-Seq Desktop Sequencer with at least 90% of targeted regions covered to a depth of at least 20 fold by PathWest Laboratories (Western Australia). Sequence alignment was performed via Illumina's BWA enrichment utilizing BWA-MEM and variant calling was performed using GATK. Variant were filtered and

analysed in Bench Lab NGS (Cartagenia) with particular focus on genes implicated in craniosynostosis and microcephaly plus *KAT6A, GNB1. GRIN1. DOCK7, ADSL* and *KANSL1*.

Individual 4, 5 and 6

Singleton WES was performed by the NATA-accredited Victorian Clinical Genetics Services Clinical Exome Laboratory (Melbourne, Australia).

Individual 7

Singleton WES was performed in cooperation with The Yale Center for Genome Analysis (Connecticut, USA) on DNA extracted from blood using Oragene kit. Array capture was done using Roche/Nimblegen SeqCap EZ Exome Library v2.0 and sequencing was performed on Illumina HiSeq platform. The mean coverage of the exome was 120X with 97-98% exome covered at least eight times and 94-95% exome was covered greater than 20X. The resulting genomic sequence was analysed for the presence of likely pathogenic variants in the customised list of genes that are often associated with development delay and ASD. Variants were filtered for the relevance to human disease and on their population frequency using proprietary exome database and curated on the basis of their likelihood to affect protein function and information in disease specific databases. Further, interpretation in context with the clinical information was performed by clinical molecular geneticist certified through American Board of Medical Genetics and Genomics. Variant is previously reported to be *de novo* in (Kennedy et al., 2019)

1.3. Supplementary figure captions

Figure S1: Pedigrees of families with RTT/RTT-like features and *KAT6A* variants along with Sanger sequencing traces.

Family pedigrees in all seven families with affected individuals with *KAT6A* variants with Sanger sequencing traces for families 1, 2 and 3 showing *de novo* status of *KAT6A* variants. Singleton exome sequencing was performed for individuals 4 and 5 in the VCGS NATA accredited laboratory but parents declined to have Sanger sequencing confirmation. *KAT6A* variants in individuals 6 and 7 have been previously reported as *de novo* (Kennedy et al., 2019). Circles = females, squares = males and diamond = unknown gender. Half-darkened symbol represents heterozygous change, arrow represents the affected individual. Sanger sequencing traces are shown alongside in the pedigree.

Figure S2: Multiple sequence alignment showing evolutionary conservation of the region affected by *KAT6A* variants in the RTT/ RTT-like individuals.

Sequences were aligned using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) by using RefSeq of *H. sapiens* (NP_006757.2), *M musculus* (NP_001074618.1), *R norvegicus* (NP_001094040.1), *M mulatta* (XP_028708463.1) and *B taurus* (XP_024842204.1). * (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a . (period) indicates conservation between residues with weakly similar properties. The amino acids affected by the variation are highlighted: p.(Arg1129*) in individual 1 (red), p.(Glu1274*) in individual 2 (yellow), p.(Lys1134Argfs*14) in individual 3 (cyan), p.(Ser1126Phefs*8) in individual 4 (green), p.(Val1211*) in individual 5 (teal) and p.(Glu1419Trpfs*12) in individual 6 (pink) and p.(Glu1221*) in individual 7 (black). Variants affect well conserved acidic domain (788-1414 aa) and/ or Ser/Met domain (1414-2004 aa).

1.4. Supplementary tables

Table S1: Additional clinical features in RTT/ RTT-like individuals

	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
Neurological							
Gross motor development delay	+; severe	+	+	+; severe	+; severe	+; severe	+; severe
Microcephaly	-	+; congenital	-	+; congenital	+; congenital	-	+
Intellectual disability	+; severe	+; severe	+; severe	+; severe	+; severe	+; severe	+; severe[@]
Seizures	+	-	+	-	+	-	-
Impaired social/emotional development	+; poor eye contact	+; poor social interaction, inconsolable crying	+; inconsolable crying	+; severe delay	+; severe delay	+	+
Behavioural problems	+	+	+	-	+; irritability and mania	+	+
Hyperventilation	+	-	-	-	+; when excited	-	-
Sleep disturbance	+; severe (apnoea, hyperventilation)	+; severe (bruxism)	+	+; sleep initiation and maintenance	+; sleep initiation and maintenance	+	+; sleep apnoea
Speech							
Speech difficulties	+; severe (one word)	+; severe (non-verbal)	+; severe (non-verbal)	+; severe (non-verbal)	+; severe (non-verbal)	+; severe	+; severe (non-verbal)
Gastrointestinal problems							
Feeding difficulties	+	+	+	+; bulbar palsy, PEG feeds	+; bulbar palsy, PEG feeds	+	+; NGT; supraglottoplasty
Gastroesophageal reflux	+	+	+	+	+	+	+
Constipation	+; severe	+; severe	+; severe	+	+	+	+; severe
Recurrent bowel obstructions	+	-	-	-	-	-	-
Intestinal malrotation	+	-	-	-	-	-	-
Gastrostomy	-	-	-	+	+	-	-
Anal stenosis	-	-	-	-	-	-	-
Duodenal rupture	-	-	-	-	-	-	-
Cardiac problems							
Congenital heart defect	-	-	-	-	+	+; PDA, PFO, pulmonary artery stenosis	+; PDA, PFO, leaky mitral valve
Atrial septal defects	+	-	-	-	+	-	+
Ventricular septal defects	+	-	-	-	-	-	-
Corrective surgery	+; Septal defects	-	-	-	-	+; PFO	-
Haematological and immunological abnormalities							

	Individual 1 (II:3) (Family 1)	**Individual 2 (II:1) (Family 2)**	**Individual 3 (II:2) (Family 3)**	**Individual 4 (II:2) (Family 4)**	**Individual 5 (II:2) (Family 5)**	**Individual 6 (II:1) (Family 6)**	**Individual 7 (II:1) (Family 7)**
Neutropenia	-	-	-	-	+	-	-
Frequent lung infections	-	-	+	-	+	-	-
Frequent herpes simplex infections	-	-	-	-	+	-	-
Frequent staphylococcus infections	-	-	-	-	+	+	-
Frequent urinary tract infections	-	+	+; six episodes; (had mild renal reflux & circumcision to improve frequency	-	-	-	-
Impetigo (skin infection)	-	-	-	-	-	+	-
Other infections	-	-	+; conjunctivitis	-	+; recurrent mouth ulcers & fevers; responds to prednisolone	+; tonsillitis	-
Otitis media	-	-	+; three episodes	-	+	-	-
Facial features							
Broad nasal tip	-	+	+	-	-	+	+
Thin upper lip	+	+	+	+	+	+	+
Ear morphology	Normal	Normal	Normal	Low set and post rotated	Normal	Normal	Low set
Bitemporal narrowing	-	+	+; mild	-	-	+	-
Prominent nasal bridge		+	+	+	+	+	-
Philtrum defects	+; long	+; short	-	+; smooth	-	-	-
High mouth palate	+	+; narrow	-	+	+	-	-
teeth abnormalities	+; wide & sparse	+; bruxism (early childhood)	+; wide & sparse	+; enamel hypoplasia	+; enamel hypoplasia	+; wide & sparse	-
Other facial feature	Very fair skin color distinct from family; gets bluish easily		Blue eyes and blonde hair (compared to parents who have brown eyes and dark hair)		Very fair skin color distinct from family	-	Recessed jaw, narrow forehead, narrow mouth
Brain abnormalities							
MRI abnormalities	-	+	-	-	+	-	-
Thin corpus callosum	-	-	-	-	-	-	-
Delayed myelination	-	+	-	-	+	-	-
Craniosynostosis	-	-	+	-	-	-	+

	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
other brain abnormalities	-	-	-		Arnold-Chiari type I malformation	-	-
Ocular problems							
Strabismus	-	-	-	+	+	-	-
Epicanthal fold	-	-	-	+	+	+	-
Optic nerve atrophy	-	-	-	-	-	-	-
Ptosis	-	-	-	-	-	-	-
Myopia	+	-	-	+	-	-	-
Other	Poor eye contact	-	Cortical vision impairment; improved slowly		Cortical visual impairment. Ocular albinism		Cortical visual impairment
Genitourinary issues							
Hydronephrosis	-	-	+; very mild	-	-	-	-
Cryptorchidism	-	-	+	NA	+	NA	-
Hypospadias	-	-	-	NA	-	NA	-
Musculoskeletal issues							
Hypotonia	-	+; infancy	+	-	+	+	+
Hypertonia	-	+; limb rigidity	-	+; four limb motor disorder with hypertonia and hyperreflexia with bulbar dysfunction	-	-	+; in infancy
Brachydactyly	-	-	-; tapered fingers	+	-; long fingers	-	-
Dystonia	-	+	+; especially in legs	-	-	-	-
Scoliosis/ kyphosis	+; scoliosis	+; kyphosis	+; mild scoliosis	+; kyphosis[#]	+; kyphosis and scoliosis[#]	-	+; mild scoliosis
Other			May have choreiform movements & dyskinesia when unwell. bilateral single palmar creases. saggital craniosynostosis treated with craniotomy and omega springs. Had	Anhi~~y~~drosis, joint hypermobility 9/9. patellar stabilisations	Broad knuckles, trichorrhexis nodosa and trichoptilosis, 5. Persistent foetal fat pads		

	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
			bilateral undescended testes with right inguinal hernia.				

=No surgical intervention needed; @ = Not formally diagnosed yet; PFO = Patent Foramen Ovale

Table S2: ACMG classification of *KAT6A* variants in RTT/ RTT-like individuals

Criteria	Individual 1 (II:3) (Family 1)	Individual 2 (II:1) (Family 2)	Individual 3 (II:2) (Family 3)	Individual 4 (II:2) (Family 4)	Individual 5 (II:2) (Family 5)	Individual 6 (II:1) (Family 6)	Individual 7 (II:1) (Family 7)
Null variant (nonsense, frameshift, canonical +/−1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where loss of function (LOF) is a known mechanism of disease **(PVS1)**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Same amino acid change as a previously established pathogenic variant regardless of nucleotide change **(PS1)**	Yes	No	No	No	No	Yes	Yes
De novo (paternity and maternity confirmed) **(PS2)**	Yes	Yes	Yes	No	No	Yes	Yes
Located in a mutational hot spot and/or critical and well-established functional domain (e.g. active site of an enzyme) without benign variation **(PM1)**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Absent in population database **(PM2)**	Yes	Yes	Yes	Yes	Yes	Yes	Yes
De novo (paternity and maternity not confirmed) **(PM6)**	No	No	No	Yes	Yes	No	No
PVS	1	1	1	1	1	1	1
PS1 to PS4	2	1	1	0	0	2	2
PM1 to PM6	2	2	2	3	3	2	2
Classification	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic	Pathogenic

2. Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

3. Conflict of interest

All authors declare that they have no conflict of interest.

4. Author contribution

Author	Contribution	ORCiD	Email
SK	Design, Data curation, Formal analysis, Project administration, Methodology, Roles/Writing - original draft	0000-0002-0689-2965	simran.kaur@mcri.edu.au
NVB	Supervision, analysis, writing – review & editing	0000-0002-6768-3665	nicole.vanbergen@mcri.edu.au
BBZ	Clinician for individual 1, writing – review & editing	0000-0001-6347-2089	Bruria.BenZeev@sheba.health.gov.il
EL	Research scientist for individual 2, writing – review & editing	0000-0001-8486-8461	emanuela.leonardi@unipd.it
TYT	Clinician for individual 6, writing – review & editing	0000-0001-8455-7778	tiong.tan@vcgs.org.au
DC	Clinician for individual 4 and 5, writing – review & editing	0000-0001-6303-6471	david.coman@health.qld.gov.au
BK	Clinician for individual 3, writing – review & editing	0000-0001-9115-9962	Benjamin.Kamien@health.wa.gov.au
SMW	Support in dysmorphic feature assessment for individuals, writing – review & editing	0000-0001-7611-634X	sue.white@vcgs.org.au
MSJ	Resources for individual 7, writing – review & editing	0000-0003-4723-937X	miya.stjohn@mcri.edu.au
DP	ACMG classification		dean.phelan@vcgs.org.au
KR	ACMG classification	0000-0001-9207-5453	kristin.rigbye@vcgs.org.au
SCL	ACMG classification	0000-0003-0312-9937	chern.lim@vcgs.org.au
MCT	ACMG classification	0000-0002-1275-4446	michelle.torres@vcgs.org.au
MM	ACMG classification	0000-0001-5035-8244	melanie.marty@vcgs.org.au
ES	ACMG classification	0000-0002-7987-3693	elena.savva96@vcgs.org.au
TZ	ACMG classification	0000-0002-3445-3898	tian.zhao@vcgs.org.au

SM	Writing – review & editing	0000-0002-0759-9306	Sean.massey@mcri.edu.au
AM	Clinician for individual 2; writing – review & editing	0000-0001-6788-0653	alessandra.murgia@unipd.it
WAG	Supervision, analysis, writing – review & editing	0000-0003-1808-0646	wendy.gold@health.nsw.gov.au
JC	Conceptualization, design, Methodology, supervision, analysis, Project administration, writing – Review & Editing, Lab head at MCRI	0000-0002-8431-0641	john.christodoulou@mcri.edu.au

5. References

Aspromonte, M. C., Bellini, M., Gasparini, A., Carraro, M., Bettella, E., Polli, R., . . . Leonardi, E. (2019). Characterization of intellectual disability and autism comorbidity through gene panel sequencing. *Hum Mutat, 40*(9), 1346-1363. doi:10.1002/humu.23822

Christodoulou, J., Grimm, A., Maher, T., & Bennetts, B. (2003). RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat, 21*(5), 466-472. doi:10.1002/humu.10194

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., & den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat, 32*(5), 557-563. doi:10.1002/humu.21438

Genomes Project, C., Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., . . . Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-74. doi:10.1038/nature15393

Kaur, S., Van Bergen, N. J., Gold, W. A., Eggers, S., Lunke, S., White, S. M., . . . Christodoulou, J. (2019). Whole exome sequencing reveals a de novo missense variant in EEF1A2 in a Rett syndrome- like patient. *Clinical case Reports, 00*, 1-7. doi:10.1002/ccr3.2511

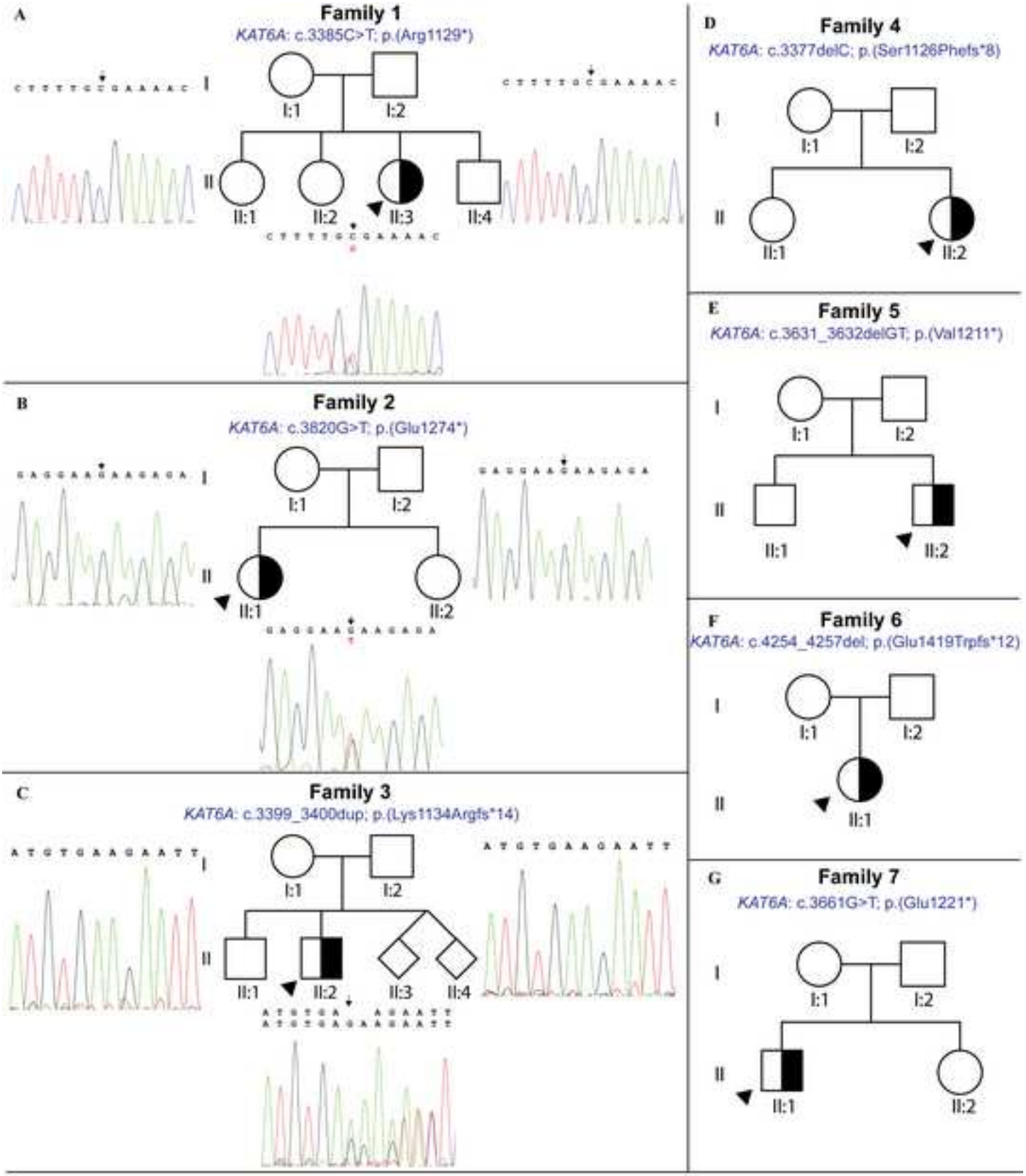
Kennedy, J., Goudie, D., Blair, E., Chandler, K., Joss, S., McKay, V., . . . Newbury-Ecob, R. (2019). KAT6A Syndrome: genotype-phenotype correlation in 76 patients with pathogenic KAT6A variants. *Genet Med, 21*(4), 850-860. doi:10.1038/s41436-018-0259-2

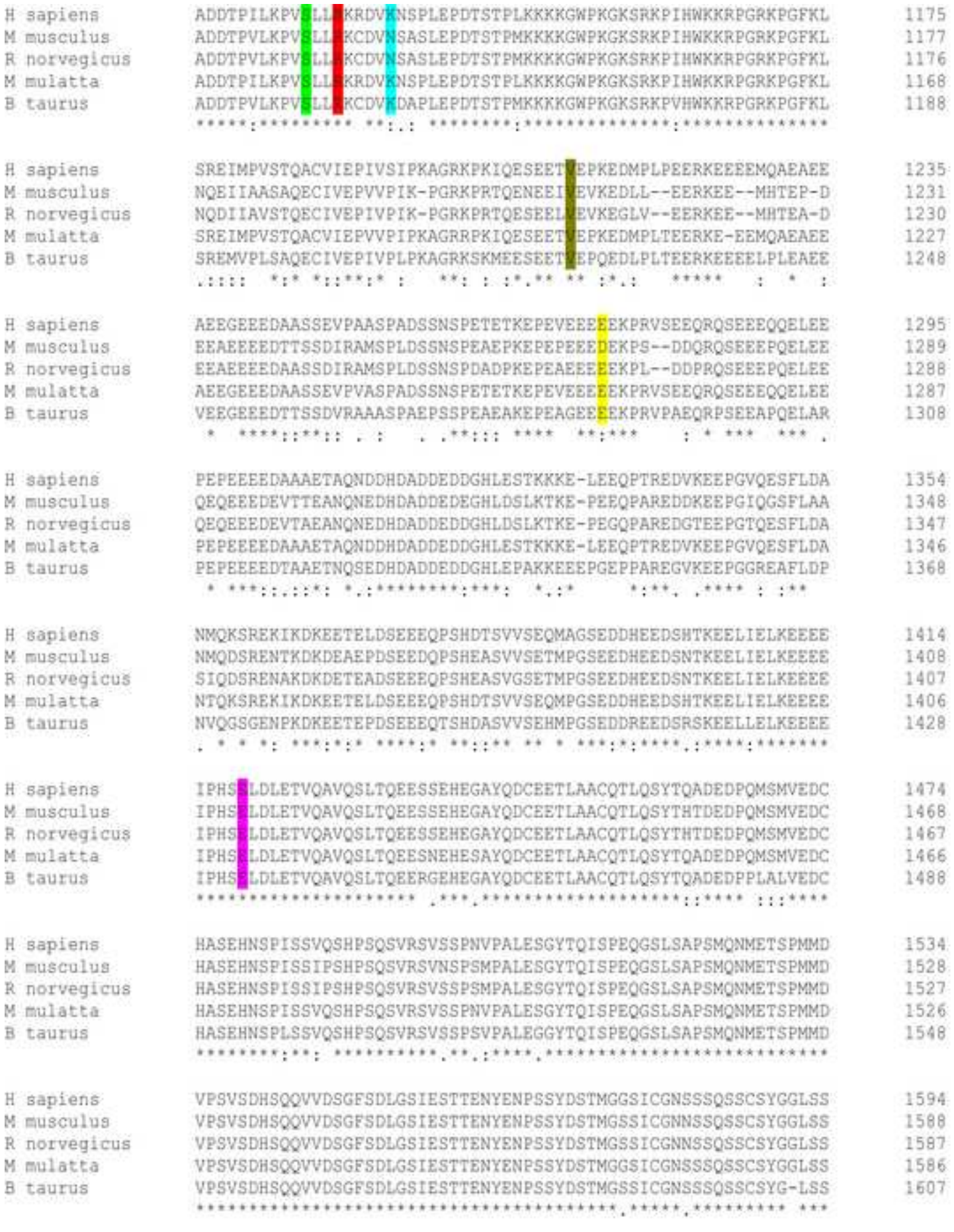
Krishnaraj, R., Ho, G., & Christodoulou, J. (2017). RettBASE: Rett syndrome database update. *Hum Mutat, 38*(8), 922-931. doi:10.1002/humu.23263

Li, Q., & Wang, K. (2017). InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet, 100*(2), 267-280. doi:10.1016/j.ajhg.2017.01.004

Sadedin, S. P., Dashnow, H., James, P. A., Bahlo, M., Bauer, D. C., Lonie, A., . . . Thorne, N. P. (2015). Cpipe: a shared variant detection pipeline designed for diagnostic settings. *Genome Med, 7*(1), 68. doi:10.1186/s13073-015-0191-x

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res, 38*(16), e164. doi:10.1093/nar/gkq603


A
Family 1
KAT6A: c.3385C>T; p.(Arg1129*)
I
II
I:1
I:2
II:1
II:2
II:3
II:4
B
Family 2
KAT6A: c.3820G>T; p.(Glu1274*)
I
II
I:1
I:2
II:1
II:2
C
Family 3
KAT6A: c.3399_3400dup; p.(Lys1134Argfs*14)
I
II
I:1
I:2
II:1
II:2
II:3
II:4
D
Family 4
KAT6A: c.3377delC; p.(Ser1126Phefs*8)
I
II
I:1
I:2
II:1
II:2
E
Family 5
KAT6A: c.3631_3632delGT; p.(Val1211*)
I
II
I:1
I:2
II:1
II:2
F
Family 6
KAT6A: c.4254_4257del; p.(Glu1419Trpfs*12)
I
II
I:1
I:2
II:1
G
Family 7
KAT6A: c.3661G>T; p.(Glu1221*)
I
II
I:1
I:2
II:1
II:2


H sapiens ADDTPILKPVSLLKRDVKNSPLEPDTSTPLKKKKGWPKGKSRKPIHWKKRPGRKPGFKL 1175
M musculus ADDTPVLKPVSLLKCDVNSASLEPDTSTPMKKKKGWPKGKSRKPIHWKKRPGRKPGFKL 1177
R norvegicus ADDTPVLKPVSLLKCDVNSASLEPDTSTPMKKKKGWPKGKSRKPIHWKKRPGRKPGFKL 1176
M mulatta ADDTPILKPVSLLKRDVKNSPLEPDTSTPLKKKKGWPKGKSRKPIHWKKRPGRKPGFKL 1168
B taurus ADDTPVLKPVSLLKCDVKDAPLEPDTSTPMKKKKGWPKGKSRKPVHWKKRPGRKPGFKL 1188
H sapiens SREIMPVSTQACVIEPIVSIPKAGRKPKIQESEETVEPKEDMPLPEERKEEEEMQAEAEE 1235
M musculus NQEIIAASAQECIVEPVVPIK-PGRKPRTQENEEIVEVKEDLL--EERKEE--MHTEP-D 1231
R norvegicus NQDIIAVSTQECIVEPIVPIK-PGRKPRTQESEELVEVKEGLV--EERKEE--MHTEA-D 1230
M mulatta SREIMPVSTQACVIEPVVPIPKAGRRPKIQESEETVEPKEDMPLTEERKE-EEMQAEAEE 1227
B taurus SREMVPLSAQECIVEPIVPLPKAGRKSKMEESEETVEPQEDLPLTEERKEEEELPLEAEE 1248
H sapiens AEEGEEEDAASSEVPAASPADSSNSPETETKEPEVEEEEEKPRVSEEQRQSEEEQQELEE 1295
M musculus EEAEEEEDTTSSDIRAMSPLDSSNSPEAEPKEPEPEEEDEKPS--DDQRQSEEEPQELEE 1289
R norvegicus EEAEEEEDAASSDIRAMSPLDSSNSPDADPKEPEAEEEEEKPL--DDPRQSEEEPQELEE 1288
M mulatta AEEGEEEDAASSEVPVASPADSSNSPETETKEPEVEEEEEKPRVSEEQRQSEEEQQELEE 1287
B taurus VEEGEEEDTTSSDVRAAASPAEPSSPEAEAKEPEAGEEEEKPRVPAEQRPSEEAPQELAR 1308
H sapiens PEPEEEEDAAAETAQNDDHDADDEDDGHLESTKKKE-LEEQPTREDVKEEPGVQESFLDA 1354
M musculus QEQEEEDEVTTEANQNEDHDADDEDEGHLDSLKTKE-PEEQPAREDDKEEPGIQGSFLAA 1348
R norvegicus QEQEEEDEVTAEANQNEDHDADDEDDGHLDSLKTKE-PEGQPAREDGTEEPGTQESFLDA 1347
M mulatta PEPEEEEDAAAETAQNDDHDADDEDDGHLESTKKKE-LEEQPTREDVKEEPGVQESFLDA 1346
B taurus PEPEEEEDTAAETNQSEDHDADDEDDGHLEPAKKEEEPGEPPAREGVKEEPGGREAFLDP 1368
H sapiens NMQKSREKIKDKEETELDSEEEQPSHDTSVVSEQMAGSEDDHEEDSHTKEELIELKEEEE 1414
M musculus NMQDSRENTKDKDEAEPDSEEDQPSHEASVVSETMPGSEEDHEEDSNTKEELIELKEEEE 1408
R norvegicus SIQDSRENAKDKDETEADSEEEQPSHEASVGSETMPGSEEDHEEDSNTKEELIELKEEEE 1407
M mulatta NTQKSREKIKDKEETELDSEEEQPSHDTSVVSEQMPGSEDDHEEDSHTKEELIELKEEEE 1406
B taurus NVQGSGENPKDKEETEPDSEEEQTSHDASVVSEHMPGSEDDREEDSRSKEELLELKEEEE 1428
H sapiens IPHSPLDLETVQAVQSLTQEESSEHEGAYQDCEETLAACQTLQSYTQADEDPQMSMVEDC 1474
M musculus IPHSPLDLETVQAVQSLTQEESSEHEGAYQDCEETLAACQTLQSYTHTDEDPQMSMVEDC 1468
R norvegicus IPHSPLDLETVQAVQSLTQEESSEHEGAYQDCEETLAACQTLQSYTHTDEDPQMSMVEDC 1467
M mulatta IPHSPLDLETVQAVQSLTQEESNEHESAYQDCEETLAACQTLQSYTQADEDPQMSMVEDC 1466
B taurus IPHSPLDLETVQAVQSLTQEERGEHEGAYQDCEETLAACQTLQSYTQADEDPPLALVEDC 1488
H sapiens HASEHNSPISSVQSHPSQSVRSVSSPNVPALESGYTQISPEQGSLSAPSMQNMETSPMMD 1534
M musculus HASEHNSPISSIPSHPSQSVRSVNSPSMPALESGYTQISPEQGSLSAPSMQNMETSPMMD 1528
R norvegicus HASEHNSPISSIPSHPSQSVRSVSSPSMPALESGYTQISPEQGSLSAPSMQNMETSPMMD 1527
M mulatta HASEHNSPISSVQSHPSQSVRSVSSPNVPALESGYTQISPEQGSLSAPSMQNMETSPMMD 1526
B taurus HASEHNSPLSSVQSHPSQSVRSVSSPSVPALEGGYTQISPEQGSLSAPSMQNMETSPMMD 1548
H sapiens VPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTMGGSICGNSSSQSSCSYGGLSS 1594
M musculus VPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTMGSSICGNNSSQSSCSYGGLSS 1588
R norvegicus VPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTMGSSICGNNSSQSSCSYGGLSS 1587
M mulatta VPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTMGGSICGNSSSQSSCSYGGLSS 1586
B taurus VPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTMGSSICGNSSSQSSCSYG-LSS 1607

Chapter 6

Chapter 6 NGS of mutation negative RTT/RTT-like individuals: Uncovering known RTT associated genes

6.1 Introduction

Pathogenic variants in *MECP2* are the most common cause of the pathology behind classic RTT syndrome, and although sequencing of the *MECP2* gene has widely been accepted as a first-tier diagnostic test for establishing the genetic diagnosis in RTT individuals, a considerable proportion of RTT individuals remain genetically undiagnosed using this approach. Further, the phenotypic spectrum of RTT-like individuals, who share similar clinical features between RTT and other NDDs is constantly evolving, thus further broadening the genetic and phenotypic spectrum of RTT/ RTT-like syndromes. Currently, due to the affordability and availability of advanced sequencing platforms, NGS methods including WGS, WES, and targeted sequencing panels are increasingly being utilized to uncover causative genes in mutation negative patients with various NDDs including RTT. This has not only led to the identification of previously missed variants in *MECP2* (Iwama et al., 2019; Sajan et al., 2017; Schonewolf-Greulich, Bisgaard, Moller, et al., 2019), but has uncovered new genetic associations in the pathogenesis of RTT.

The three most common genes identified in RTT individuals are *MECP2*, *CDKL5* and *FOXG1* (Ariani et al., 2008; Neul et al., 2010; Zhang et al., 2017). RTT has been classified as an epigenetic disorder due to the global regulatory role of MeCP2, which binds to methylated CpG dinucleotides to regulate transcriptional repression and activation, chromatin remodelling and RNA splicing (Sanfeliu et al., 2019). Evidence of RTT being an underlying epigenetic disorder is further supported by the fact that FOXG1 is a regulator of neurogenesis, and is also involved in global chromatin organisation (Manuel et al., 2011; Zaret & Mango, 2016). The evidence of common pathways between MeCP2 and CDKL5 (Carouge et al., 2010; Mari et al., 2005) further supports a causal association between defects in epigenetic regulation and RTT. Interestingly, NGS approaches performed on mutation negative RTT/ RTT-like individuals uncovering additional novel variants in genes associated with chromatin regulation such as *ACTL6B*, *BRD1*, *CHD4*, *CREB1*, *HDAC1*, *SMARCB1*, *SMC1A*, *SET* and

TRRAP, further implicating epigenetic dysregulation in the pathophysiology of RTT (Jang et al., 2015; Richardson et al., 2018; Sajan et al., 2017; S. J. C. Stevens et al., 2018). In addition, defects in genes contributing to glutamate receptor signalling (*GABRB2, GRIN2A, GRIN2B,* and *SHANK3*) have also been found to be involved in RTT (Sajan et al., 2017) and dysfunction of the protein translational machinery may be another contributory to RTT, as evidenced by our recent identification of RTT patients with pathogenic variants in *EEF1A2* (Kaur et al., 2019; Lopes et al., 2016).

In addition to finding novel disease genes in RTT individuals, NGS techniques have also aided in recognizing pathogenic variants in *MECP2* previously missed by Sanger sequencing. For example, Gilissen et al (2014) identified a deletion within exon 4 of *MECP2* through trio WGS, and retrospective analysis revealed that the primers used in the initial diagnostic *MECP2* Sanger sequencing were located within a deleted region, and so the variant was not identified using this method. In addition, this diagnosis had also been missed using Multiplex Ligation Probe Amplification (MLPA), routinely used for copy number variants, since the probes were outside the deleted region, resulting in apparently normal results (Gilissen et al., 2014). In addition, Sajan et al. identified three classic RTT individuals with pathogenic *MECP2* variants that were initially missed during routine diagnostic testing (Sajan et al., 2017).

This demonstrates that NGS is uncovering previously missed pathogenic variants in RTT-associated genes and also expanding the genetic landscape of RTT/ RTT-like syndromes. A clearer overview of additional genes implicated in novel molecular pathways will improve the genetic diagnosis for affected individuals and open up opportunities for individualized drug therapies.

6.2 Methods

Using singleton NGS techniques, a cohort of 42 mutation negative RTT/RTT-like individuals were investigated using sequencing techniques previously described in Chapter 2, to find additional genes associated with their clinical phenotype. All the identified potentially

interesting variants with predicted likely pathogenicity were then validated by Sanger sequencing of parental samples, when available, to confirm the segregation.

6.3 Results

6.3.1 Methyl CpG Binding Protein 2 (*MECP2*)

Through WES, previously reported pathogenic variants in *MECP2* were identified in seven mutation negative classic RTT probands. These variants are as listed in Table 6.1. Three RTT individuals (**JC001**, **BST003** and **BBZ006**) were found to carry the same pathogenic variant [c.397C>T; p.(Arg133Cys)] in *MECP2*. This variant has been previously reported as one of the three most recurrent missense variants in the MBD domain, along with p.(Arg106Trp) and p.(Thr158Met), and results in a milder phenotype in the previously reported individuals as compared to other missense variants affecting *MECP2* (Leonard et al., 2003).

Table 6.1: *MECP2* variants identified in seven RTT individuals

Patient code	*MECP2* variant	Location of variant
JC003	c.104_108del; p.(Lys35fs)	Exon 3
JC001	c.397C>T; p.(Arg133Cys)	Exon 4
BST003	c.397C>T; p.(Arg133Cys)	Exon 4
BBZ006	c.397C>T; p.(Arg133Cys)	Exon 4
JC006 (Sa_Ha)	c.481G>T; p.(Gly161Trp)	Exon 4
BBZ003	c.806delG; p.(Gly269Alafs*20)	Exon 4
BST026	c.1308dupT; p.(Gln437fs*)	Exon 4

Genbank transcript ID: NM_004992 (also known as *MECP2A*, includes exon 2); NCBI Protein ID: NP_004983.1

Of the seven individuals identified with *MECP2* variants, one case (**BST026**) was an 8 year old (as of 2018) genetically undiagnosed male with a classic RTT phenotype according to the Neul revised diagnostic criteria (Neul et al., 2010). The proband's key clinical features included regression followed by stabilization at 18 months of age, stereotypic hand movements, absent speech, poor hand skills, ataxia, intellectual disability and poor eye contact. Other RTT features observed in this individual were breathing problems, bruxism, impaired sleep, abnormal muscle tone, peripheral vasomotor disturbances, small & cold extremities, inappropriate

laughing and crying spells and reduced response to pain. Previous WES studies using DNA isolated from his whole blood did not identify a causative gene in this case. RTT is less commonly diagnosed in males due to its X-linked dominant inheritance pattern (Reichow et al., 2015). Moreover, males with previously identified pathogenic variants in *MECP2* usually present with an earlier onset and more severe phenotype compared to females, since all cells express the pathogenic defected *MECP2* allele (Psoni et al., 2010; Schule et al., 2008). In order to identify the causative variant, we performed targeted sequencing of DNA extracted from a muscle biopsy using our custom designed gene panel. This revealed a previously reported pathogenic variant in *MECP2* [c.1308dupT; p.(Gln437fs*)], but was mosaic (Table 6.1). To determine whether the identified mutation was mosaic across various tissues, our collaborators determined the degree of mosaicism in DNA extracted from different tissues. Variable degrees of mosaicism (percentage indicated in brackets) was present across a range of samples including DNA from blood (9%), buccal swab (16%), fibroblasts (12%) and muscle (15%). Thus, the resulting lower variant allele frequency in the blood may account for the diagnosis being missed during earlier WES studies due to a more stringent variant filtering threshold. This study suggested that analysis of *MECP2* variants in RTT males should be carried out using deep-sequencing on DNA extracted from multiple tissues. These findings have been published as a short report "*Mosaic MECP2 variants in males with classical Rett syndrome features, including stereotypical hand movements*" in Clinical genetics (this male RTT individual is listed as Case A in the publication) (Schonewolf-Greulich, Bisgaard et al., 2019) (Appendix G, Publication 1).

6.3.2 Structural Maintenance of Chromosomes 1A (*SMC1A*)

Targeted panel sequencing was performed using customized our gene panel (Appendix A, List A2) on DNA isolated from whole blood from a *MECP2* negative classic RTT female. She also had additional clinical features including epilepsy, progressive microcephaly, limb defects, congenital heart disease, renal malformations and dysmorphic facial features. Data analysis from WES identified a novel, heterozygous frameshift variant [NM_001281463.1:c.3576delA; NP_001268392.1: p.(Val1193Serfs*2)] in an X-linked gene called Structural Maintenance of Chromosomes 1A (*SMC1A*; OMIM: 300040). *In-silico* analysis predicted this variant to be

likely-pathogenic, and Sanger sequencing performed by our collaborators confirmed this variant to be *de novo*.

The *SMC1A* gene encodes one of the structural subunits of the Cohesin complex that holds sister chromatids together, and ensures proper chromosomal segregation during cell division (Ishiguro, 2019). The Cohesin complex core mainly consists of two structural protein subunits; SMC1A and SMC3, which interact with four regulators that are critical to determine its function. The interacting proteins include NIPBL (homologue of fungal Scc2 and *Drosophila* Nipped-B) for loading cohesion onto chromatin structures, ESCO2, which forms connections between sister chromatids, and PDS5A/5B proteins, which regulate the dynamic interaction between the Cohesin complex and chromatin. In addition, SMC1A, SMN3 and NIPBL play an important role in regulation of gene expression (Revenkova et al., 2009), DNA repair and maintenance (Jessberger et al., 1993; Jessberger et al., 1996).

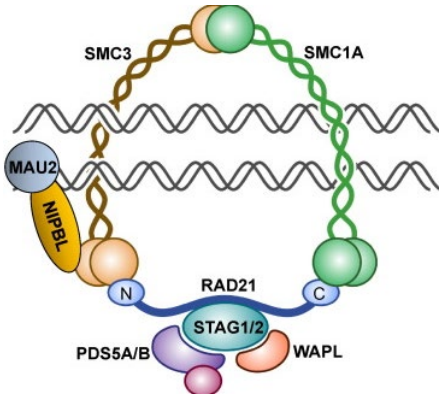


Figure 6.1: Cohesin complex structure.

SMC1A and SMC3 form the two main structural components of the cohesin complex, which further associates with regulatory proteins such as NIPBL, PDS5A/5B for regulating chromatin dynamics. Image taken from (Hill, Kim, & Waldman, 2016).

While heterozygous variants in *NIPBL* cause a form of cohesinopathies, Cornelia de Lange syndrome type 1 (CdLS1), various pathogenic variations in *SMC1A* have been identified in individuals with Cornelia de Lange syndrome 2 (CdLS2) (Mannini, Liu, Krantz, & Musio, 2010). CdLS is a group of clinically heterogeneous multisystemic developmental disorders characterised by intellectual disability, growth retardation, distal limb defects, gastrointestinal abnormalities and typical dysmorphic facial features including thick-arched eyebrows, long eye lashes, synophrys (unibrow created due to merging of two eyebrows), and downward slanting palpebral fissures (Kline et al., 2018). In addition, several individuals with RTT/ RTT-like clinical phenotypes have been genetically associated with the likely-pathogenic variants in

SMC1A; making it the fourth most common gene associated with RTT-like individuals (Gilissen et al., 2014; Huisman et al., 2017; Sajan et al., 2017). The identification of a RTT individual with this novel variant in *SMC1A* further broadens the genotypic spectrum of RTT/RTT-like individuals with *SMC1A* variants.

These findings have been published in collaboration as a part of the review article "Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature" in Clinical Genetics 95(2):221-230 (Schonewolf-Greulich et al., 2019) (Appendix G, Publication 2).

6.3.3 SH3 and multiple ankyrin repeat domains 3 (*SHANK3*)

The DNA from a RTT-like female fulfilling three main, and four supportive criteria of Neul's revised diagnostic criteria (Neul et al., 2010) and exhibiting classic autistic behaviour, was sequenced using our customized targeted gene panel. She was found to have a *de novo* heterozygous splice site variant [chr22:51,153,476G>A; NM_033517.1: c.2265+1G>A] in the SH3 and multiple ankyrin repeat domains 3 (*SHANK3*) gene.

SHANK3 encodes a ubiquitously expressed scaffolding protein that is mainly enriched in post-synaptic densities of excitatory synapses (Yi et al., 2016). Neurons generated from a heterozygous and a homozygous conditional knockout of *SHANK3* gene in human H1 embryonic stem cells resulted in abnormal neuronal morphology, dendritic arborisation defects, altered neuronal activity and impaired hyperpolarization-activated cation channels (Yi et al., 2016). The splicing variant identified in the RTT proband affected

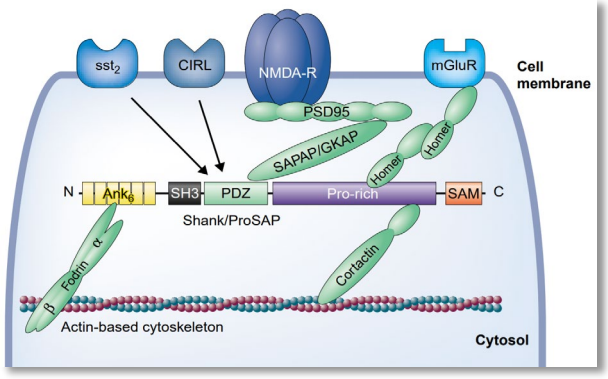


Figure 6.2: SHANK3 as a scaffolding synaptic protein

Various domains of SHANK3 interact with other distinct proteins at the synapse that are critical for neuronal structure and function. Image taken from (Hughes, 2012).

intron 19 of *SHANK3,* and was predicted to cause skipping of exon 19; ultimately resulting in a frameshift and premature stop of the SHANK3 protein (Bramswig et al., 2015). Due to unavailability of the affected RTT individual's cells, the effect of the identified variant to cause either protein truncation or NMD could not be tested. Regardless, it is likely that the truncated protein would impact the interaction of SHANK3 with other synaptic proteins including Homer and PSD95, both of which have been related with ASDs (Guang et al., 2018).

Heterozygous pathogenic variations spanning the length of *SHANK3* (Figure 6.4) have been associated with Phelan-McDermid syndrome, a rare neurodevelopmental disorder characterised by intellectual disability, global development delay, abnormal muscle tone and ASDs (De Rubeis et al., 2018). Additionally, haplo-insufficiency in *SHANK3* has been linked with approximately 1% of ASD cases, thus making it one of the leading autism candidate disease genes (De Rubeis et al., 2018). The pathogenic *de novo* heterozygous variant (c.3100del, p.Ala1034fs*44) in *SHANK3* has also been reported in a RTT-like female (Hara et al., 2015). The *de novo* heterozygous splicing variant (c.2265+1G>A) identified in our RTT-like female is one of the most recurrent single nucleotide variants in *SHANK3.* This variant and has been previously reported in an individual with severe intellectual disability, absent speech and mild facial dysmorphism (Bramswig et al., 2015; Holder & Quach, 2016; RK et al., 2017). Thus, our findings further expand the phenotypic spectrum of variants in SHANK3 to include individuals with a RTT/ RTT-like phenotype.

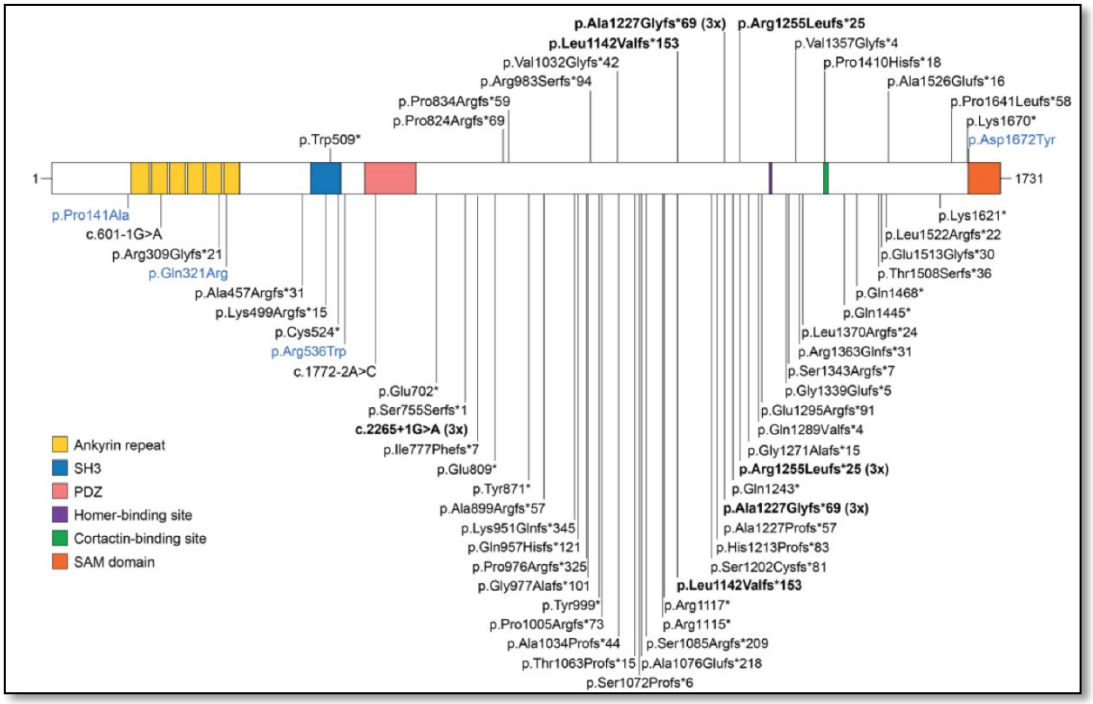


Figure 6.3: SHANK3 domain structure and mutation spectrum.

Domain structure of SHANK3 showing the key domains: Ankyrin repeats (yellow), SH3 domain (blue), PDZ domain (light red), homer-binding site (purple), cortactin-binding site (green) and SAM domain (orange). Likely pathogenic and pathogenic variants in SHANK3 reported in the literature and ClinVar. Black text denotes loss-of-function variants, blue text represents missense variants and bold text signifies recurrent variants. Image taken from (De Rubeis et al., 2018).

The finding related to the *SHANK3* variant in our affected individual has been already published in collaboration as a part of the review article "Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature" in Clinical Genetics 95(2):221-230 (Schonewolf-Greulich et al., 2019) (Appendix G, Publication 2). This article provides an overview of RTT-related genes and highlights the clinical and genetic heterogeneity of RTT/ RTT-like syndrome, and provides suggestions for additional genetic approaches to aid in diagnosing RTT/RTT-like individuals.

6.3.4 Elongation factor 1-alpha 2 (*EEF1A2*)

Using singleton WES, in which only proband DNA was screened, a heterozygous variant [NM_001958.3:c.271G>A; p.(Asp91Asn)] affecting a conserved residue in Elongation factor 1-alpha 2 (*EEF1A2*; OMIM: 602959) (Figure 6.1) was identified in a 7.5 year old RTT-like female with global developmental delay, stereotypic hand movements, abnormal breathing, limited communicative abilities, and inappropriate outbursts of laughter. Sanger sequencing of parental DNA samples confirmed this variant to be *de novo.*

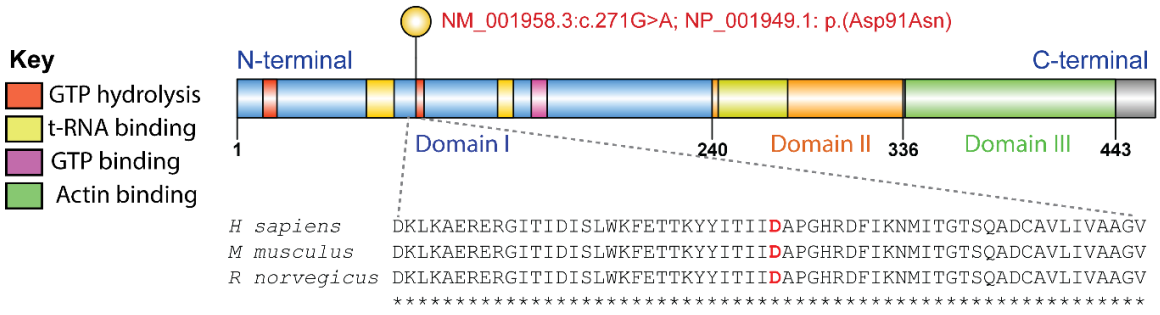


Figure 6.4: EEF1A2 domain structure and location of patient variant.

The domain structure of EEF1A2 including three sub-domains Domain I, Domain II and Domain III containing sequences for GTP hydrolysis (red), t-RNA binding (yellow), GTP binding (purple), and actin binding (green). The variant affects the first residue of the critical GTP hydrolysis domain and the position is well conserved between different species including human (NP_001949.1), mouse (NP_031932.1) and rat (NP_036792.2) as shown in the multiple sequence alignment by using CLUSTAL Omega (version 1.2.4; https://www.ebi.ac.uk/Tools/msa/clustalo/).

These findings have been published as a short report "Whole exome sequencing reveals a *de novo* missense variant in *EEF1A2* in a Rett syndrome-like patient" in Clinical Case reports **7**(12): 2476-2482 (Kaur et al., 2019). The full article is included on next page.

Received: 31 May 2019 | Revised: 19 September 2019 | Accepted: 30 September 2019

DOI: 10.1002/ccr3.2511

CASE REPORT

Clinical Case Reports Open Access WILEY

Whole exome sequencing reveals a de novo missense variant in *EEF1A2* in a Rett syndrome-like patient

Simranpreet Kaur[1,2] | **Nicole J. Van Bergen**[1,2] | **Wendy Anne Gold**[3,4] | **Stefanie Eggers**[5,6] | **Sebastian Lunke**[5,6] | **Susan M. White**[2,6] | **Carolyn Ellaway**[4,7] | **John Christodoulou**[1,2,4,6]

[1]Brain and Mitochondrial Research Group, Murdoch Children's Research Institute, Parkville, Vic., Australia

[2]Department of Paediatrics, University of Melbourne, Parkville, Vic., Australia

[3]Molecular Neurobiology Lab, Kids Research, Westmead Children's Hospital, Westmead, NSW, Australia

[4]Disciplines of Genetic Medicine and Child and Adolescent Health, Sydney Medical School, University of Sydney, NSW, Australia

[5]Translational Genomics Unit, Murdoch Children's Research Institute, Parkville, Vic., Australia

[6]Victorian Clinical Genetics Services, Murdoch Children's Research Institute, Parkville, Vic., Australia

[7]Genetic Metabolic Disorders Service, Sydney Children's Hospital Network, Sydney, NSW, Australia

Correspondence
John Christodoulou, Brain and Mitochondrial Research Group, Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Vic. 3052 Australia.
Email: john.christodoulou@mcri.edu.au

Funding information
This research was supported by internal funding support to JC.

Abstract

Using whole exome sequencing, we found a pathogenic variant in the EEF1A2 gene in a patient with a Rett syndrome-like (RTT-like) phenotype, further confirming the association between EEF1A2 and Rett syndrome RTT and RTT-like phenotypes.

KEYWORDS

EEF1A2, elongation factor-1, intellectual disability, mutation, Rett syndrome

1 | INTRODUCTION

Rett syndrome (RTT; OMIM 312750) is a severe, rare X-linked neurodevelopmental disorder that primarily affects females and is notable for its progressive nature.[1,2] There are two major clinical categories of RTT syndrome: classical (or typical) RTT and atypical (or variant) RTT. Patients may also be classified as RTT-like who do not meet the diagnostic criteria for a clinical diagnosis of classical or atypical RTT.[3] In classical RTT, after a period of normal growth and development, patients exhibit distinct stages of disease. During stage I (6 - 18 months of age), RTT females do not meet expected milestones due to developmental stagnation. Between 1 and 4 years (stage II) a rapid developmental regression phase features loss of acquired skills (eg, hand use and speech) with impaired social contact and patients develop stereotypic hand movements (wringing and washing). Microcephaly may become evident. Between 4 and 10 years (stage III) the clinical phenotype may stabilize, but affected individuals may develop breathing irregularities, seizures, and intense eye gaze. Beyond 10 years (stage IV), patients lose ambulation and experience muscle weakness, rigidity, spasticity, dystonia, and scoliosis; however, the hand stereotypic movements become less intense while eye contact and communication remain intact. Atypical RTT patients share many key of the supportive clinical features seen in classical RTT but do not have all of the main criteria.[3] Almost all classical RTT patients (~98%) have pathogenic variants in the X-linked Methyl CpG Protein 2 gene (*MECP2*),[4] while only 60% of atypical RTT patients have pathogenic variants in *MECP2*.[5]

Although *MECP2* is the major causative gene for classical and atypical RTT, mutations in other genes including cyclin-dependent kinase-like 5 (*CDKL5*), Forkhead box protein G1 (*FOXG1*), myocyte-specific enhancer factor 2C (*MEF2C*), and transcription factor 4 (*TCF4*) have also been found in patients with clinical profiles overlapping with RTT.[6]

Here, we report a RTT-like patient who was negative for mutations in the more common genes associated with RTT. Previous genetic analyses failed to identify the causative gene. We identified a de novo heterozygous variant previously associated with intellectual disability but not RTT in the eukaryotic translation elongation factor 1 alpha 2 (*EEF1A2*: chr20 [q13.33]: 62,119,365-62,130,668:11,304bp) gene.

2 | CLINICAL REPORT

This study was approved by the Sydney Children's Hospitals Network Human Research Ethics Committee with written consent obtained from the parents. The proband was born at 39 weeks gestation following a pregnancy complicated by gestational diabetes managed with diet and normal delivery to nonconsanguineous parents. She flipped into a breech position at around 37 weeks and was delivered by cesarean section. Her Apgar scores were 9 at one and 9 at 5 minutes. Her birthweight was 2.75 kg (10th percentile), length was 50.5 cm (50th percentile), and head circumference was 31.5 cm (<2nd percentile). Blood sugar levels were monitored in the newborn period and were normal. She smiled at 6 weeks and was described as a sleepy baby. As she was breech, she had a routine hip scan at 6 weeks of age, which showed mild hip displacement, with a follow-up scan at 4 months of age revealing hip dysplasia, which was managed in a Pavlik harness for 4 weeks. The hip dysplasia subsequently resolved, as evidenced by a normal pelvic X-ray at 2 years of age (Figure 1). At 4 months of age, she was noted to have increased tone, scissoring of her legs and stared at her hands. At 6 months, she was thought to have a seizure. An EEG showed mild slowing for age but no epileptiform activity, while a head ultrasound was normal. At 14 months, she was seen by a Paediatric Neurologist who noted that she had small cold, hands and feet, was bringing her hands to her mouth, was babbling but not saying any words. The electroencephalogram and brain magnetic resonance imaging (MRI) were normal at this time. By 2 years and 3 months, her weight was 8.64 kg (2nd percentile), length was 72 cm (2nd percentile), and head circumference was 43 cm (<1st percentile). No facial dysmorphism was noticed by two experienced clinical geneticists. She was unable to roll or sit unsupported, articulate distinct words and had limited hand function. She had started to exhibit midline hand clasping movements and some hand-to-mouth movements. Her hands and feet were cold and sometimes mottled. She had feeding difficulties, constipation, and bruxism and was found to have central hypothyroidism. Seizures developed three months later. She had abnormal muscle tone and now exhibited an abnormal breathing pattern. She also had limited communicative abilities, inappropriate outbursts of laughter and a diminished response to pain. She could not sit unsupported and had poor head control. When last reviewed at 5 years of age, her weight was 14.3 kg (1st percentile), height approximately 100 cm (1st percentile), and head circumference was 46 cm (1st percentile). She continued to have episodes of breath-holding and deep sighing respirations but no hyperventilation. She continued to grind her teeth, although not as frequently as previously. Her eye contact had improved. She had developed some additional stereotypic hand movements in the form of hand-to-mouth movements, sucking her knuckles, midline hand clasping, and tapping or rubbing her stomach. She had central hypotonia with increased tone of her limbs. Cardiovascular and respiratory examinations were normal. She also had a postural upper thoracic kyphosis and was able to stand with assistance.

3 | GENETIC ANALYSIS

As part of the genetic diagnosis for the patient with RTT-like features, sequencing of the *MECP2* gene at the Molecular Genetics Department, Children's Hospital at Westmead, did not identify any sequence variation and no deletion or duplication was detected by multiplex ligation-dependent probe amplification studies. Subsequent genomic analysis of an epileptic encephalopathy gene set using the TruSight One

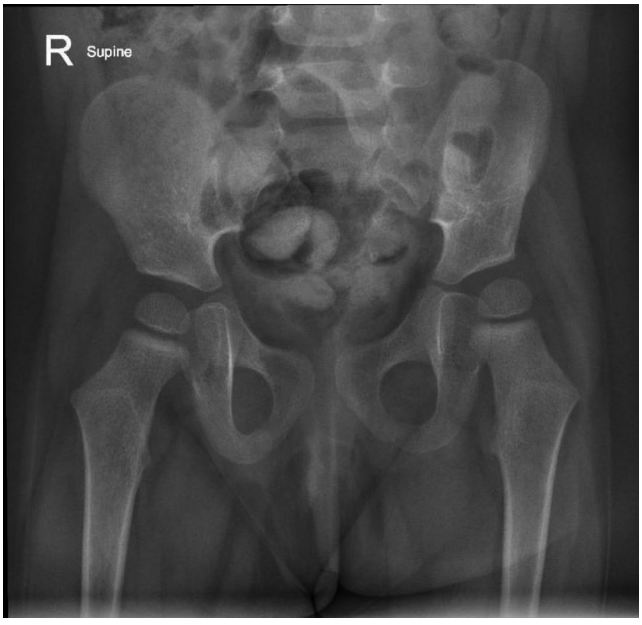


FIGURE 1 Pelvic X-ray at 2 y of age showing resolution of the hip dysplasia

sequencing panel at The Children's Hospital at Westmead, NSW, Australia, was also negative, although the coverage for some genes including *FOXG1* and *CDKL5* was not 100%. Singleton WES was therefore undertaken in order to try to identify the causative gene. WES SureSelectQXT Clinical Research Exome libraries were prepared and loaded onto a NextSeq 500 sequencer (Illumina; NextSeq control Software v2.1.031) and 2 × 150 bp paired-end sequencing was performed at the Translational Genomics Unit, Victorian Clinical Genetic Services (VCGS). Samples passed sequencing QC with >89.7% bases with at least Q30, and mean coverage of at least 100-fold. Raw sequencing data were converted to FASTQ format using Illumina's bcl2fastq2 converter (v2.17.1.14). Data were processed using Cpipe (http://cpipeline.org/),[7] in order to generate annotated variant calls within the target region (coding exons ± 2bp), via alignment to the reference genome (GRCh37). Variants were annotated against all gene transcripts, with reporting of variants against the HGNC recommended transcript (according to HGVS nomenclature). Classification of variants was based on ACMG guidelines.[8] A total of 25 679 variants were found via Cpipe analysis. All variants were uploaded onto Leiden Open (source) Variation Database (LOVD) prior to further analysis.[9] First, the RTT associated genes were screened to examine any previously missed variants.[10,11] We then used a precurated phenotype-specific gene list (https://www.vcgs.org.au/sites/default/files/media/TGW024_genelist_V3_0.pdf) provided by the VCGS for targeted analysis of known intellectual disability (ID) associated genes. This list consists of a total of 1,064 genes implicated in syndromic and nonsyndromic ID.

An unknown singleton inheritance filter was used to filter out variants with low quality (<100), genomic mean allelic frequency (>5%), and intronic splice region variants, untranslated region (UTR) variants, synonymous as well as low impact variants. The remaining 550 variants were further prioritized based upon the impact of the variant in the order of early stop gain, frameshift and missense variants. Pathogenicity of variants was evaluated using in silico tools including SIFT (http://sift.jcvi.org/), MutationTaster (www.mutationtaster.org/), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), CADD, and Grantham scores.[12,13] The final shortlisted variants were then classified using The American College of Medical Genetics and Genomics (ACMG) guidelines [8] to interpret the consequence of the sequence variants. A literature search was conducted on each of the filtered variants to identify their implications in known disorders. In addition, 3-dimensional molecular modeling was performed using the already published protein crystal structure through the automatic variant analysis server HOPE (http://www.cmbi.ru.nl/hope/).[14]

Using this filtering process, variants were prioritized based upon variant characteristics, in silico predictions of pathogenicity and expression in brain. We identified a heterozygous missense variant in *EEF1A2* (chr20: g.62127262C > T; NM_001958.3: c.271G > A; p.(Asp91Asn)) that has been previously reported [15] (Figure 2A). This variant was confirmed by Sanger sequencing DNA from child's parents to be de novo (Figure 2B). Based on multiple sequence alignment of the protein sequence using CLUSTAL O (1.2.4) (https://www.ebi.ac.uk/Tools/msa/clustalo/), the position of this aspartic acid is very highly conserved between different species including human (NP_001949.1), mouse (NP_031932.1), and rat (NP_036792.2) (Figure 3A). This variant is predicted to create an amino acid change from aspartic acid (Asp) to asparagine (Asn) at amino acid position 91 (Figure 3B). This variant was absent in population databases including Exome Aggregation Consortium (http://exac.broadinstitute.org) and gnomAD (http://gnomad.broadinstitute.org/). In silico analysis predicted the variant to be damaging (SIFT, score 0.001), deleterious (PROVEN, score −3.85), and disease causing (MutationTaster) with a CADD Phred-like score of 19.88 and Grantham score of 23. Based on the current in silico analysis, this variant has been classified as likely pathogenic based on ACMG criteria.

4 | DISCUSSION

In this study, we have identified a de novo heterozygous variant in EEF1A2 (NM_001958.3: c.271G > A; p.(Asp91Asn)) in a patient with a RTT-like phenotype. The identified NM_001958.3 (EEF1A2): c.271G > A substitution variant is predicted to change Asp to Asn at amino acid position 91, NP_001949.1 (EEF1A2): p.(Asp91Asn) affecting an evolutionarily conserved site in exon 3, and is situated within the guanosine triphosphate (GTP) binding domain. Heterozygous de novo mutations in EEF1A2 have previously been associated with neurodevelopmental disorders including epilepsy, autism, and severe intellectual disability.[15-24] Moreover, three siblings with dilated cardiomyopathy, global development delay, failure to thrive, epilepsy and ultimately death in early childhood have been reported to carry a homozygous mutation in EEF1A2 (NM_001958: c.1164C > T; p.Pro333Leu).[25] The variant identified in our case has been previously reported as a de novo heterozygous variant in a 14-year-old female patient with, epilepsy, intellectual disability, hypotonia, cold peripheries, lack of speech, and inability to walk independently.[15] Our patient shares some clinical features with the patient reported by Lam et al,[15] including cold peripheries, seizures, hypotonia, and spine deformity. In addition, our patient also exhibits distinct features typically seen in RTT, including stereotypic hand movements, teeth grinding, limited speech, feeding difficulties, breath-holding, inappropriate outbursts of laughter, a diminished

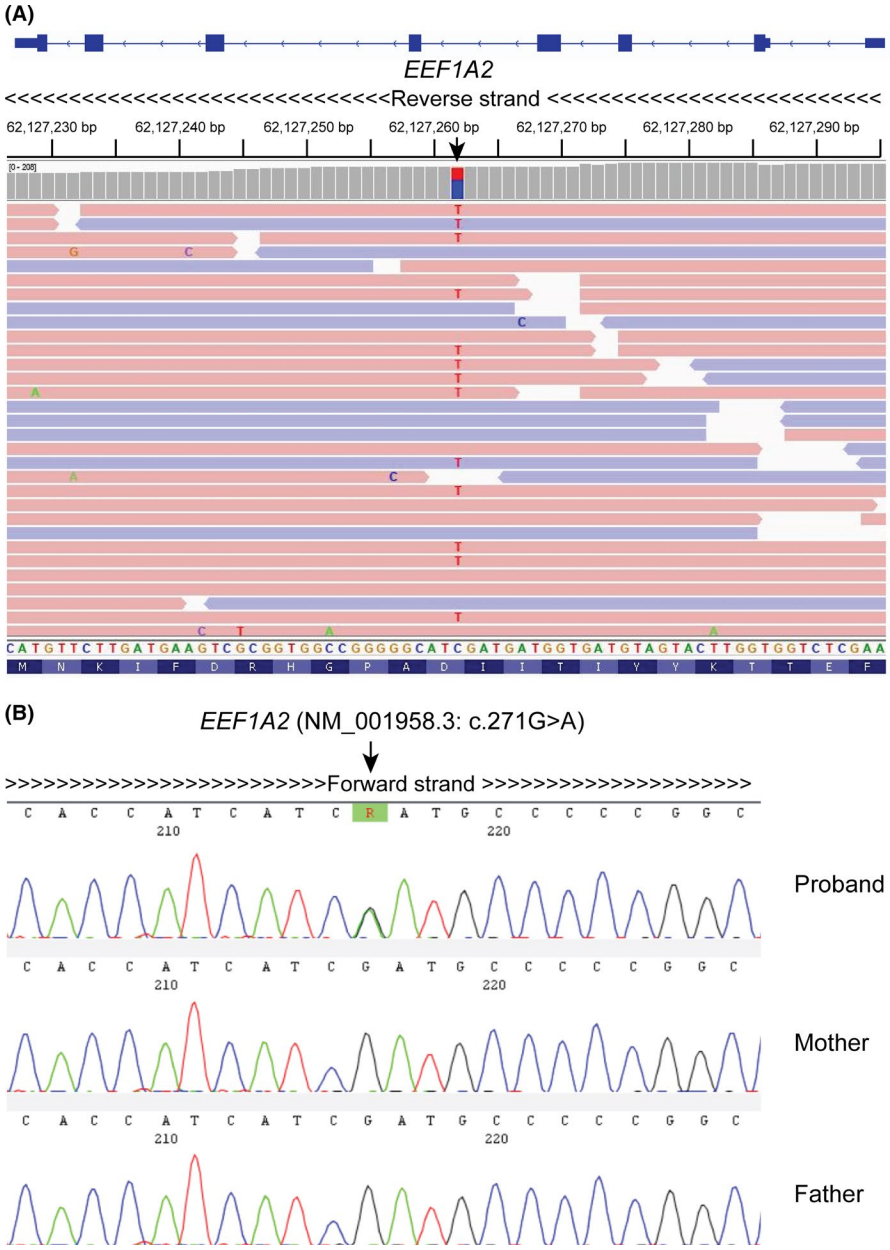


FIGURE 2 A, Schematic showing the Integrative Genomics Viewer reads (reverse strand) of whole exome sequencing encompassing the heterozygous variant (NM_001958.3 (*EEF1A2*): c.271G > A; p.(Asp91Asn)) reported in the proband. Note that *EEF1A2* is a reverse strand gene and hence the variant is shown as C > T change on the IGV reads. B, Sanger sequencing chromatogram (forward strand) showing a confirmed heterozygous variant for *EEF1A2* variant in the proband with both parents being homozygous wild type

response to pain, and no facial dysmorphism. This highlights that a variable neurodevelopmental phenotype may be associated with *EEF1A2* variants, with the same variant resulting in overlapping neurological presentations. Interestingly, a single case with a de novo heterozygous variant NM_001958: c.274G > A, p.(Ala92Thr) has also been reported in case of a 6-year-old RTT-like female patient with stereotypic hand movements, loss of spoken language but no regression. She had seizures at 1 month of life and her development was significantly delayed, with delayed onset of speech (3 years) and ability to walk unaided (4 years). The patient also had hand stereotypies, bruxism, and crying spells when awake, sleep problems, hyperpnoea and apnea, and poor eye contact and was classified as RTT-like.[18]

Three-dimensional molecular modeling of the EEFA12 protein (Protein Data Bank code: 4C0S) performed using HOPE [14] predicted a change in the charge of the wild-type (Asp; negative) to the mutant residue (Asn; neutral). Moreover, the wild-type residue forms a hydrogen bond with Ala92 and Asn101, and a salt bridge with Arg67 [26] which are predicted to be disrupted by the mutated residue. The difference in charge is likely to disturb ionic interactions created by the wild-type residue and thus interfere with the conformational changes of the EEFA1A protein.

The eEF1A2 protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis. During this process, eEF1A2 transits between an active GTP bound state and an inactive GDP bound state thus classified as Translational GTPase (trGTPases), a family of proteins in which GTPase activity of the protein is stimulated by the large ribosomal subunit. The other subunit, eEF1Bαβγ, then catalyses the nucleotide exchange of GDP for GTP in order to re-activate the eEF1A

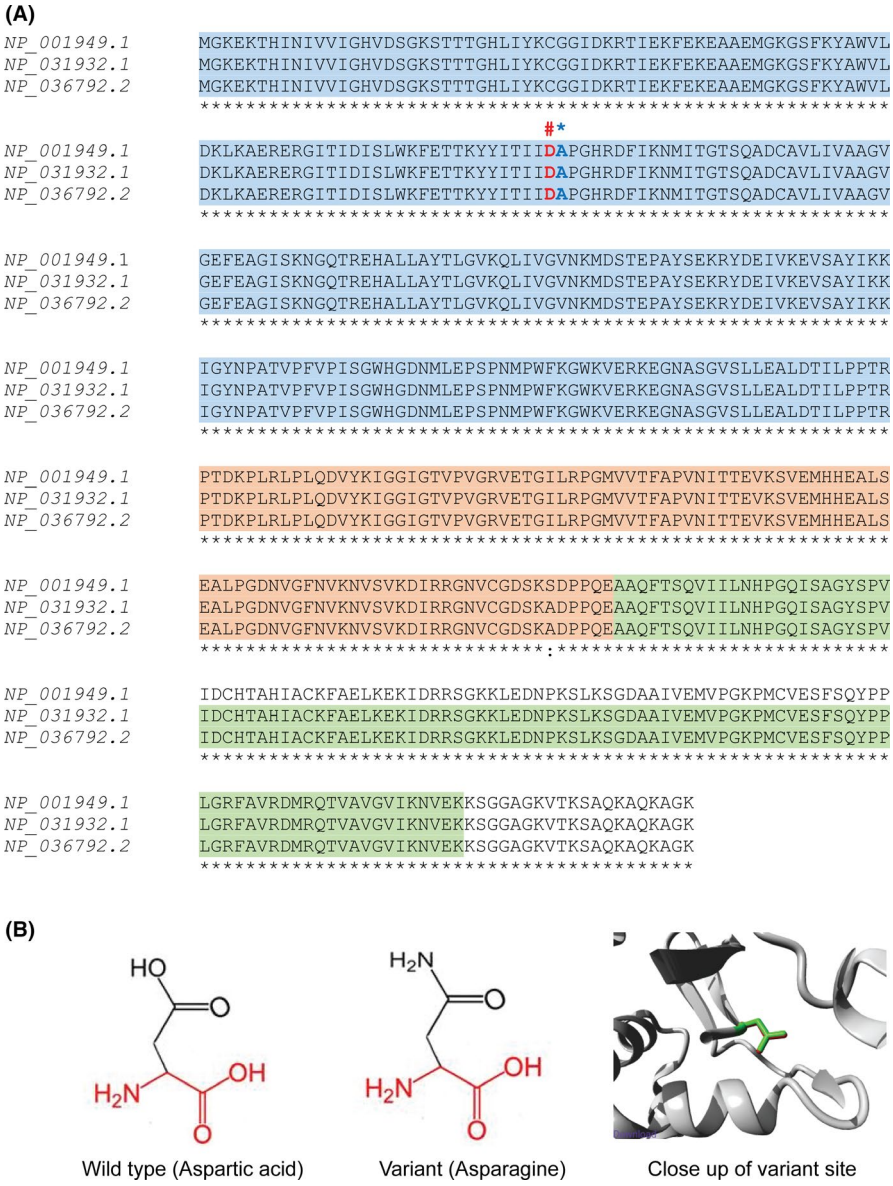


FIGURE 3 A, Evolutionary sequence conservation of EEF1A2. Schematics represents the multiple sequence alignment of EEF1A2 protein from human (NP_001949.1), mouse (NP_031932.1), and rat (NP_036792.2) performed using CLUSTAL O (1.2.4) (https://www.ebi.ac.uk/Tools/msa/clustalo/).The three structural domains are highlighted as domain I (residue 1-240, highlighted blue), domain II (residue 241-336, highlighted orange), and domain III (residue 337-443, highlighted green). Domain I contains a helix that associates with the GTP and GDP. Domain I and domain II contain the eEF1B complex binding site for GTP/GDP exchange. Domain II and domain III harbor the aa-tRNA binding site. Domain III contains an actin-binding domain. The affected amino acid residue Asp91 (D) identified in our patient is marked (#) and the mutated residue Ala92 (A) described in Lopes, Barbosa et al (2016) is marked (*). B, Schematic structures of the wild type and variant amino acids with the close up of the variant site modeled with HOPE website. The protein is colored in gray, the side chain of the wild type and variant is green and red respectively. The backbone, which is the same for each amino acid, is colored red, the side chain, unique for each amino acid, is colored black

complex for the next cycle of hydrolysis. Because of its functional role, binding, and hydrolysis of guanine (G) nucleotides, the G-binding domain is a highly conserved between different species (Figure 3A). The amino acid Asp91 resides in this highly conserved domain and is stabilized by other amino acids when GDP is bound to the complex. This reaction is compromised during GDP displacement by the amino acid change we identified in this patient and consequently is predicted to affect the catalytic function of the protein. Because of the critical role of de novo protein synthesis machinery at the synaptic terminal, and the enrichment of eEF1 complex at the postsynaptic density region, the optimal functioning of the complex becomes critical for proper neuronal development. Any perturbation to this complex (in particular eEF1A2) is likely to result in neurodevelopmental dysfunction, as evidenced by affected individuals exhibiting

varying clinical features including autism, epilepsy, and intellectual disability.[16-18,25,27] Therefore, our in silico analysis supports this variant to be disease causing for this RTT-like patient.

5 | CONCLUSIONS

We have identified a de novo variant in *EEF1A2* in a patient with a Rett-like phenotype that confirms the association between *EEF1A2* and a RTT-like phenotype. This study also further emphasizes the clinical utility of singleton WES to identify causative genes in *MECP2* negative patients with RTT and RTT-like phenotypes, and we suggest that analysis of *EEF1A2* should be included in the curation of genomic sequencing data from individuals with a RTT or RTT-like clinical picture.

ACKNOWLEDGMENTS

We thank all the family members for taking part in this study. The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

SK: involved in design, experimental, analysis, and writing. NJVB and WAG: involved in design, analysis, and writing. SE and SL: involved in sequencing and writing. SMW and CJE: clinical input and writing. JC: design, analysis, writing, and overall oversight.

ORCID

Wendy Anne Gold https://orcid.org/0000-0003-1808-0646

John Christodoulou https://orcid.org/0000-0002-8431-0641

REFERENCES

1. Rett A. On a unusual brain atrophy syndrome in hyperammonemia in childhood. *Wien Med Wochenschr*. 1966;116(37):723-726.
2. Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann Neurol*. 1983;14(4):471-479.
3. Neul JL, Kaufmann WE, Glaze DG, et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol*. 2010;68(6):944-950.
4. Percy AK, Lane J, Annese F, Warren H, Skinner SA, Neul JL. When Rett syndrome is due to genes other than MECP2. *Transl Sci Rare Dis*. 2018;3(1):49-53.
5. Dolce A, Ben-Zeev B, Naidu S, Kossoff EH. Rett syndrome and epilepsy: an update for child neurologists. *Pediatr Neurol*. 2013;48(5):337-345.
6. Armani R, Archer H, Clarke A, et al. Transcription factor 4 and myocyte enhancer factor 2C mutations are not common causes of Rett syndrome. *Am J Med Genet A*. 2012;158A(4):713-719.
7. Sadedin SP, Dashnow H, James PA, et al. Cpipe: a shared variant detection pipeline designed for diagnostic settings. *Genome Med*. 2015;7(1):68.
8. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
9. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat*. 2011;32(5):557-563.
10. Christodoulou J, Grimm A, Maher T, Bennetts B. RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat*. 2003;21(5):466-472.
11. Krishnaraj R, Ho G, Christodoulou J. RettBASE: Rett syndrome database update. *Hum Mutat*. 2017;38(8):922-931.
12. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
13. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073-1081.
14. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*. 2010;11:548.
15. Lam W, Millichap JJ, Soares DC, et al. Novel de novo EEF1A2 missense mutations causing epilepsy and intellectual disability. *Mol Genet Genomic Med*. 2016;4(4):465-474.
16. Veeramah KR, Johnstone L, Karafet TM, et al. Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia*. 2013;54(7):1270-1281.
17. Nakajima J, Okamoto N, Tohyama J, et al. De novo EEF1A2 mutations in patients with characteristic facial features, intellectual disability, autistic behaviors and epilepsy. *Clin Genet*. 2015;87(4):356-361.
18. Lopes F, Barbosa M, Ameur A, et al. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet*. 2016;53(3):190-199.
19. Lelieveld SH, Reijnders MRF, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci*. 2016;19(9):1194-1196.
20. Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216-221.
21. Inui T, Kobayashi S, Ashikari Y, et al. Two cases of early-onset myoclonic seizures with continuous parietal delta activity caused by EEF1A2 mutations. *Brain Dev*. 2016;38(5):520-524.

22. Epi K, Bellows ST, Berkovic SF, et al. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol*. 2017;16(2):135-143.
23. de Ligt J, Willemsen MH, van Bon BWM, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012;367(20):1921-1929.
24. Ohba C, Haginoya K, Osaka H, et al. De novo KIF1A mutations cause intellectual deficit, cerebellar atrophy, lower limb spasticity and visual disturbance. *J Hum Genet*. 2015;60(12):739-742.
25. Cao S, Smith LL, Padilla-Lopez SR, et al. Homozygous EEF1A2 mutation causes dilated cardiomyopathy, failure to thrive, global developmental delay, epilepsy and early death. *Hum Mol Genet*. 2017;26(18):3545-3552.
26. Crepin T, Shalak VF, Yaremchuk AD, et al. Mammalian translation elongation factor eEF1A2: X-ray structure and new features of GDP/GTP exchange mechanism in higher eukaryotes. *Nucleic Acids Res*. 2014;42(20):12939-12948.
27. de Kovel CG, Brilstra EH, van Kempen MJA, et al. Targeted sequencing of 351 candidate genes for epileptic encephalopathy in a large cohort of patients. *Mol Genet Genomic Med*. 2016;4(5):568-580.

How to cite this article: Kaur S, Van Bergen NJ, Gold WA, et al. Whole exome sequencing reveals a de novo missense variant in *EEF1A2* in a Rett syndrome-like patient. *Clin Case Rep*. 2019;7:2476–2482. https://doi.org/10.1002/ccr3.2511

6.4 Discussion

NGS analysis of a cohort of 42 genetically undiagnosed RTT individuals resulted in the identification of pathogenic *MECP2* variants in seven individuals (~17%) that were previously missed during routine diagnostic *MECP2* genetic testing. Interestingly, deep-sequencing using a custom targeted gene panel revealed a mosaic *MECP2* variant in DNA extracted from muscle in a classic RTT male for whom previous WES testing on blood DNA was negative (Schonewolf-Greulich, Bisgaard, Duno, et al., 2019). Additionally, three genes previously associated with RTT (*EEF1A2*, *SMC1A* and *SHANK3*) were also identified in another three RTT/RTT-like individuals (~7%) (Kaur et al., 2019; Schonewolf-Greulich et al., 2019). Our findings have provided additional evidence in the involvement of errant proteins encoded by these genes in the pathophysiology of RTT.

Overall, this study has uncovered previously missed pathogenic variants in *MECP2*, provided a better diagnosis of mosaic *MECP2* individuals and further reinforced the significance of several additional genes in mutation negative RTT/RTT-like individuals.

Chapter 7

Chapter 7 NGS of mutation negative RTT/RTT-like individuals: Uncovering new RTT associated genes

7.1 Introduction

Traditionally, along with *MECP2*, only a handful of other genes had been implicated in the aetiology of RTT individuals, including Forkhead box protein G1 (*FOXG1*) and Cyclin-dependent kinase-like 5 (*CDKL5*), both of which are now regarded as distinct syndromes (Fehr et al., 2013; Wong et al., 2019). With the feasibility of NGS techniques and advanced bioinformatics analytical tools, an increasing number of genes linked with NDDs are now being associated with individuals exhibiting RTT/ RTT-like features (Cogliati et al., 2019; Henriksen et al., 2018; Iwama et al., 2019; Schonewolf-Greulich et al., 2019; Vidal et al., 2019; Yoo et al., 2017). Thus, the genetic landscape of RTT and RTT-like individuals is constantly evolving, and with a better understanding the underlying genetic cause, a diagnosis can end the often-prolonged diagnostic "odyssey" endured by patients and their families. Furthermore, uncovering new genes may uncover hidden aspects of key biological pathways behind the devastating syndrome for which there is still no cure.

In this study, we further investigated the 42 mutation-negative RTT/ RTT-like individuals described in Chapter 6 with NGS techniques in a bid to uncover new genes that have not been previously recognized to cause RTT.

7.2 Methods

7.2.1 Patient cohort, genetic analysis and variant curation

All methods used in this study were approved by the respective institutional human research ethics committee as outlined in Chapter 2. The clinical features of the affected individuals were assessed according to the Neul revised diagnostic criteria for RTT (Chapter 1; Table 1.1) (Neul et al., 2010). WGS (individual 1), WES (individuals 1, 2 and 3) and whole genome SNP microarray (individual 2) were performed by the NATA-accredited VCGS laboratory using

methods as outlined in Chapter 2 (Section 2.2). NGS was performed on DNA extracted from blood, and segregation of identified variants was confirmed using Sanger sequencing of proband, and when available, parental DNA. Variant filtering and curation was performed using methods outlined in Chapter 2 (Section 2.2.1, Section 2.2.2 and Section 2.2.3).

7.2.2 Cell culture, Immunoblotting and Immunofluorescence

Primary fibroblast cultures from individual 1 and individual 2 were cultured using methods outlined in Chapter 2 (Section 2.3.7). For immunoblotting, whole cell extracts were extracted using RIPA buffer (Chapter 2; Section 2.4.1) and denaturing SDS-PAGE was performed using 4-15% Tris-glycine-SDS polyacrylamide gels (Biorad) (Chapter 2; Section 2.4.2). The membranes were probed with primary antibodies specific to the C-terminus of CHD8 (Cell Signaling; Cat. No. 11891; 1:2000) and anti-GAPDH (Sigma Aldrich; Cat. No. G9545; 1:10,000) as a loading control for samples from individual 1. For individual 2, two different antibodies against SET (Abcam; Cat. No. ab1183 and GeneTex; Cat. No. GTX113834; 1:500) and NUP188 (Bethyl Laboratories Inc; Cat no. A302-322A-M; 1:1000) were used in which case β-Actin (Santa Cruz Biotechnology; Cat. No. sc-47778; 1:1000) was used as a loading control. Appropriate secondary anti-mouse or anti-rabbit horseradish peroxidase-conjugated antibodies (GE Healthcare) were used to detect the primary antibodies using enhanced chemiluminescence reagents (GE Healthcare) and Amersham Hyperfilm. ImageJ was used to quantify the intensity of observed protein bands in the linear range. The data was normalized to the intensity of GAPDH or β-Actin.

Fibroblasts from individual 2 were plated on sterile glass coverslips, allowed to adhere overnight, then fixed with 4% paraformaldehyde for 20 minutes at room temperature, washed three times with 1X PBS, permeablized with 0.5% Triton X-100 and blocked using 1% BSA in 1X PBS with 0.1% Tween-20 (PBST) for 1 hour at room temperature. Cells were then probed with mab414 antibody (Abcam; Cat. No. ab24609; 1:500) in 1% BSA in PBST overnight at 4ºC, washed and then counter-stained with polyclonal goat-anti-mouse IgG (H+L) -Alexa 488 (Invitrogen) secondary antibody at 1:2000 along with Hoechst (Thermo Fisher Scinetific; Cat. no. H3570; 1µg/ml) in 1% BSA for 1 hour at room temperature. After washing three times with 1X PBS, coverslips were then mounted on a clean microscope slide using ProLong Gold mountant (Thermo Fisher Scinetific; Cat. No. P36934) and sealed using varnish. The

fluorescent signal was captured using confocal microscope with Airyscan and the intensity of mab414 staining at the perimeter of the nucleus was quantified using ImageJ.

7.2.3 RNA fluorescence *in situ* hybridization (RNA-FISH)

To examine the mRNA export defects in individual 2 with potential pathogenic variants in *NUP188*, RNA-FISH was performed using methods as previously described (Wickramasinghe et al., 2010). Briefly, control and individual's fibroblasts were grown to 80% confluency followed by fixation with 4% paraformaldehyde for 5 minutes at room temperature, followed by ice-cold 100% ethanol at 4ºC for overnight. Fixed cells were stained with oligo (dT) primer labelled with Cy3 (Sigma) and 4'-6'- Diamidino-2-phenylindole dihydrochloride (DAPI) (Sigma; Cat. no. D8417). Imaging was performed using the confocal LSM780 microscope and the ratio of average fluorescence intensity in nucleus versus cytoplasm was measured and compared through ImageJ and GraphPad Prism.

7.2.4 Constructs, oocyte preparation/ injection and electrophysiology recording

The GABA-B receptor is a hetero-dimer consisting of two subunits GABA-B1 and GABA-B2, where GABA-B2 couples with G protein for receptor function (Padgett & Slesinger, 2010). G-protein signalling pathways through GABA-B receptors involve one of the three effector proteins: G protein-activated inwardly-rectifying K+ channels (GIRK) harbouring two subunits (GIRK1 and GIRK2) characterized by slow activation and deactivation kinetics, voltage-gated Ca2+ channels and adenylyl cyclase (Figure 7.1) (Padgett & Slesinger, 2010). In addition, several auxiliary subunits mainly comprised of potassium channel tetramerization-domain (KCTD)

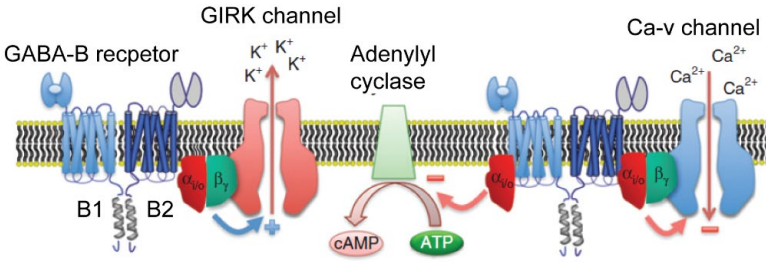


Figure 7.1: GABA-B receptor based G-protein signalling pathway.

Activation of the GABA-B receptor results in dissociation of G-protein (Gi/o) from Gα and Gβγ dimer. While Gαi/o inhibits adenylyl cyclase, the Gβγ dimer modulates voltage gated Ca^{2+} channels (Ca-v) and G protein-gated inwardly rectifying K^{+} channels (GIRK), thus resulting in neuronal inhibition. Image taken from (Padgett & Slesinger, 2010).

family members, KCTD8, KCTD12, KCTD12b and KCTD16, binds to the C-terminal tail of GABA-B2, thereby further modulating the receptor signalling kinetics by promoting the KCTD-subtype dependent receptor desensitization and agonist potency (Schwenk et al., 2010).

Previous studies of the GABA-B receptor function have been performed via heterologous expression of G-protein signalling pathway proteins including GABA-B subunits (B1 and B2), G protein-activated inwardly-rectifying K+ channels (GIRK) subunits along with KCTD proteins using *Xenopus* oocytes (M. Li et al., 2017; Vorobiov, Levin, Lotan, & Dascal, 1998), or mammalian cell lines such as CHO cells (Schwenk et al., 2010). For proof of principle studies to determine the pathogenicity of Potassium Channel Tetramerization Domain Containing 16 (*KCTD16*) variant identified in our RTT-like individual, we used a *Xenopus* oocyte model system to record changes in the electrophysiology of GABA receptor in response to the variant KCTD16. The constructs expressing human GABA-B receptor subunits [*GABBR1B* (NM_021903) and *GABBR2* (NM_005458)], effector ion channels [*GIRK1* (NM_002239), *GIRK2* (NM_002240)] and auxiliary subunit (*KCTD16*) were cloned into a *Xenopus* oocyte high expression vector (Liman, Tytgat, & Hess, 1992). The starting constructs were synthesised by Genscript (Piscataway, NJ). Along with wild-type and the variant in individual 2 [p.(Ser313Thr)], a frequent population variant p.(Asp160Asn) was also included for comparative analysis since this variant is predicted to not affect function. All open reading frames were located between the restriction sites *HindIII* and *BamHI* of the pGEMHE-mcs vector. All cDNA transcripts were transcribed into cRNA *in vitro* (mMessage mMachine, Ambion) and cRNA integrity was assessed spectrophotometrically.

Stage V or VI oocytes were surgically removed from mature *Xenopus laevis* as previously described (Petrou, Ugur, Drummond, Singer, & Walsh, 1997) and then stored in ND96 solution containing 96mM NaCl, 2mM KCl, 1.8mM $CaCl_2$, 1mM $MgCl_2$, 5mM HEPES, adjusted to pH 7.5 with NaOH. The cRNA mixture were delivered into oocytes via microinjection. The ratio of the cRNA mixture was 1:1:1:1:15 for GIRK1:GIRK2:GABBR1B:GABBR2: KCTD16, and a total of 4-19 ng of cRNA was injected into each oocyte. The chosen cRNA ratio was based upon the amount of KCTD12, another member of KCTD family, required for receptor desensitization in *Xenopus* model system (M. Li et al., 2017; Schwenk et al., 2010). The total injection volume was 100 nl per oocyte. Oocytes were kept at 17-18 ˚C for 3-4 days before

recording. At the time point of recording, the ND96 storage solution was replaced with an external recording solution containing 52 mM NaCl, 40 mM KCl, 1.8 mM $CaCl_2$, 1 mM $MgCl_2$ and 5 mM HEPES, pH 7.5 with KOH, 20-30 minutes before recording in order to stabilize the basal GIRK currents (Vorobiov et al., 1998). Two electrode voltage clamp (TEVC) recording was performed at -50mV using a TEC-10X voltage clamp (NPI Electronics). Voltage clamp control, data acquisition and superperfusion valve switching was under software control using pClamp version 8–10 software (Molecular Devices), and data were acquired using a Digidata 1332a (Molecular Devices) acquisition unit. Sampling frequency was 50 kHz, and the recording temperature was 18–20 °C. Oocytes were impaled with electrodes that contained 1.5M K-acetate and 0.5M KCl, and were held at -50 mV. After 40 seconds of baseline recording, baclofen (GABA-B receptor agonist that induces GIRK currents; 1 μM; previously optimised by our collaborators) was applied for 60 seconds, followed by 2 minutes wash out, and readings were taken on a continuous basis. Electrophysiological data was analysed using AxoGraph (AxoGraph) and presented as mean ± SEM. The changes in GABA-B receptor kinetics were measured using various parameters including K^+ raw and peak current (μA). Receptor desensitization was evaluated as the 20-80% rise time (sec). The weighted time constant tau (msec) was calculated to evaluate the decay response of GABA-B receptor. In order to compare any changes in the rate of onset of receptor response, the parameter 50% onset was analysed between wild type and variant *KCTD16* (M. Li et al., 2017; Schwenk et al., 2010).

7.3 Results

In this study, through NGS, we have identified a number of variants in genes not previously associated with RTT. Our WGS revealed a novel heterozygous nonsense variant in Chromodomain helicase DNA-binding protein 8 (*CHD8*) [c.5017C>T; p.(Arg1673*)] which led to a significant reduction in CHD8 protein levels in fibroblasts from an affected atypical RTT individual (individual 1). Molecular karyotyping (or SNP microarray) performed in a RTT-like case (individual 2) identified a *de novo* deletion at chromosome 9q34.11 (hg19:131,455,942-131,743,585) encompassing 13 genes. The most significant finding in the deleted region is the loss of the 3' end of SET Nuclear Proto-Oncogene (*SET*) and the 5' region of the Nucleoporin-188 kDa (*NUP188*). This deletion resulted in the truncation of the C-terminal region of SET, containing the nucleosome assembly protein (NAP) domain that is

critical for histone binding and nucleosome assembly (Fujii-Nakata et al., 1992; Park & Luger, 2006). In addition, the deletion resulted in the loss of critical N-terminal domain of nucleoporin protein NUP188, which is important for maintaining the structure of whole nucleoporin complex (NPC). Fibroblasts from Individual 2 showed reduced NUP188 protein levels and non-specific bands using an anti-SET antibody. Preliminary RNA-FISH studies showed a trend towards enlarged nuclei in fibroblasts from individual 2 implying reduced RNA export; however our preliminary RNA-FISH studies did not show any significant difference. Immunofluorescence studies using the mab414 antibody that demonstrates the integrity of the NPC complex, showed no structural difference in the intact nucleoporin complex. Interestingly, in individual 3 with a clinical diagnosis of classic RTT, a homozygous variant [c.937T>A; p.(Ser313Thr)] was identified in a novel candidate disease gene, *KCTD16* that has not been previously reported to be associated with any human disease. KCTD16 is one of the auxiliary subunits of GABA-B receptor. Although *in-silico* analysis predicted this variant to be likely disease causing, our preliminary electrophysiology studies in *Xenopus* oocyte co-expressing GABA and GIRK subunits along with either wild type or variant KCTD16 did not show a significant difference in GABA-B receptor desensitization in response to baclofen, which is a GABA-B agonist.

7.3.1 Clinical summary

The detailed clinical features and previous investigations performed on individuals 1, 2 and 3 are outlined in Appendix H.

Individual 1 was a 19 years old female, eldest of two daughters born to healthy non-consanguineous parents. She had severe intellectual disability, ataxia with broad-based gait, absent speech and poor eye contact. She also exhibited hypotonia, intermittent hyperventilation, poor sleep initiation and inappropriate crying/screaming spells. According to the Neul revised clinical diagnostic criteria she was diagnosed with atypical RTT (three main and five supportive criteria) (Neul et al., 2010) (Table 7.1). Genetic testing for *MECP2* did not reveal any pathogenic variants.

Individual 2 was a 24 years and 4 months old female, the eldest of three daughters born to non-consanguineous parents. She exhibited development delay, growth retardation, intellectual disability, severe speech delay with only a few words, diminished hand skills, stereotypic hand wringing/ mouthing movements, deteriorating gait, slow regression of gross and fine motor skills without recovery or stabilization, significant constipation and involuntary head writhing movements. She also had small/ cold hands and feet, peripheral vasomotor disturbances, impaired sleep, diminished response to pain, intense eye pointing and inappropriate laughing and screaming spells. In addition, she exhibited severe behavioural problems currently managed by medication. Moreover, she had dysmorphic facial features including a broad nasal tip, chin dimple and had a past history of a divergent squint (Figure 7.2). Based on the clinical information, she was classified as RTT-like as she fulfilled three main and eight supportive criteria for RTT according to the Neul revised diagnostic criteria (Neul et al., 2010) (Table 7.1). No pathogenic variants were identified in the common RTT associated genes *MECP2*, *CDKL5* and *FOXG1*.

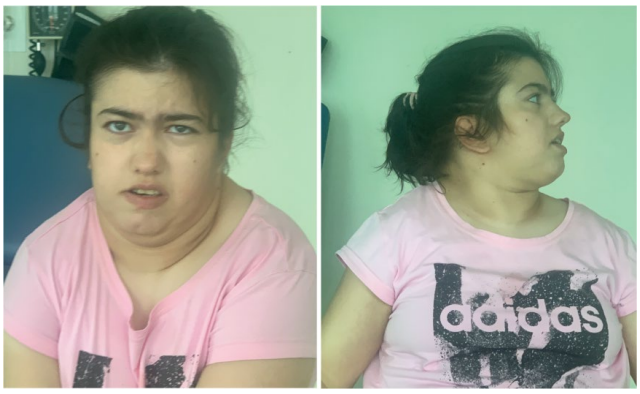


Figure 7.2: Clinical photographs of individual 2.

These photographs were taken at 24 years and 4 months of age. Her dysmorphic facial features include a broad nasal tip and chin dimple.

Individual 3 was a female born to unrelated Ashkenazi Jewish parents and had six developmentally normal siblings. She suffered from gait abnormalities, loss of hand use, stereotypic hand movements, seizures, microcephaly, mild apnoeic episodes, and absence seizures currently managed well with medication. She also exhibited bruxism, severe constipation and no distinct dysmorphic facial features. According to Neul revised diagnostic criteria (Neul et al., 2010), she was classified as classic RTT as she fulfilled all of the main and exclusion criteria (Table 7.1).

Table 7.1: Neul's diagnostic criteria assessment for all individuals.

Category	Individual 1 Family 1 (II:1)	Individual 2 Family 2 (II:1)	Individual 3 Family 3 (II:1)
Age (as of Feb 2020)	19 years and 3 months	24 years and 4 months	Unknown
Gender	Female	Female	Female
Main criteria			
Partial/ complete loss of acquired purposeful hand skills	No	Yes	Yes
Partial/ complete loss of acquired spoken language	Yes	No	Yes
Gait abnormalities: Impaired/absence of ability	Yes	Yes	Yes
Stereotypic hand movements†	Yes	Yes	Yes
Exclusion criteria			
Brain injury‡	No	No	No
Grossly abnormal psychomotor development (first 6 months)	No	Yes	No
Supportive criteria for atypical RTT			
Breathing disturbances when awake	Yes	No	Yes
Bruxism when awake	Yes	Yes	Yes
Impaired sleep pattern	Yes	Yes	Yes
Abnormal muscle tone	Yes	No	Yes
Peripheral vasomotor disturbances	No	Yes	Yes
Scoliosis/kyphosis	No	No	Yes
Growth retardation	No	Yes	Yes
Small cold hands and feet	No	Yes	Yes
Inappropriate laughing/screaming spells	Yes	Yes	No
Diminished response to pain	No	Yes	Yes
Intense eye communication - ''eye pointing''	No	Yes	Yes
Required for classic (typical) RTT			
Regression followed by recovery/ stabilization	Yes	No	Yes
All main criteria and all exclusion criteria	No		Yes
Required for atypical (variant) RTT			
Regression followed by recovery/ stabilization	Yes	No	-
At least 2 of the 4 main criteria	Yes (3/4)	Yes (3/4)	-
5 out of 11 supportive criteria	Yes (5/11)	Yes (8/11)	-
Classification	**Atypical RTT**	**RTT-like**	**Classic RTT**

†: hand wringing/squeezing, clapping/ tapping, mouthing and washing/rubbing automatisms; ‡: secondary to trauma (peri- or postnatally), neurometabolic disease, or severe infection that causes neurological problems

7.3.2 Variant curation and functional validation

The detailed variant curation of all variants is summarised in Table 7.2.

Individual 1: CHD8 variant

Using singleton WGS, no pathogenic variants in RTT-related genes (*MECP2*, *CDKL5* and *FOXG1*) were identified, however a novel heterozygous non-sense variant in *CHD8* [chr14:21866016G>A; NM_001170629.1:c.5017C>T; NP_001164100.1:p.(Arg1673*)] was identified (Figure 7.3, A). This variant is absent in gnomAD and results in the loss of a highly

conserved C-terminal Brahma and Kismet (BRK) domain involved in transcriptional control (Barnard, Pomaville, & O'Roak, 2015) (Figure 7.3, A; Appendix H, Figure H1). Several individuals with intellectual disability, ASDs and ADHD harboring likely pathogenic protein truncating variants downstream of the individual 1 variant site have been already reported (Bernier et al., 2014). Sanger sequencing confirmed the variant in DNA extracted from the affected individual's fibroblasts and blood. The variant is absent in the mother (Figure 7.3, B), however DNA from the father was unavailable to confirm the segregation. Other variants of uncertain significance (VUS) were identified in two other genes (*IMPDH2* and *KIF23*) as described in Supplementary information (Appendix H).

To determine CHD8 protein levels, we performed immunoblotting on whole cell lysates prepared from primary skin fibroblasts from individual 1. Western blot showed markedly reduced CHD8 protein levels as compared to control fibroblasts (Figure 7.3, C). The anti-CHD8 antibody was targeted to the C-terminal region of protein, downstream of the variant identified in individual 1 deleting the BRK domain.

Individual 2: chromosome 9q34.11 deletion including the SET and NUP188 genes

A *de novo* deletion was identified using trio whole genome SNP microarray in individual 2 at chromosome 9q34.11 (hg19:131,455,942-131,743,585) encompasses 13 genes including the 3' region of *SET*, all of *PKN3*, *ZDHHC12*, *LOC100506100*, *ZER1*, *TBC1D13*, *ENDOG*, *SPOUT1*, *KYAT1*, *LRRC8A*, *PHYHD1*, *DOLK* and the 5' region of *NUP188* (Figure 7.4, A; Appendix H, Table H1 and Figure H2). Singleton WES was performed to identify other potential likely causative disease genes. No pathogenic variants in any known RTT genes; *MECP2*, *CDKL5* and *FOXG1* were identified. Variants of uncertain significance (VUS) (Appendix H, Figure H2) were identified through WES, but they were subsequently confirmed to be present in either mother or father by Sanger sequencing, asserting their non-pathogenicity. The deletion at chromosome 9q34.11 identified by molecular karyotyping was also independently confirmed by WES in individual 2 (Figure 7.4, B) and was found to affect the evolutionary well-conserved residues in the NAP domain of SET (Figure 7.4, C). We then specifically focused on searching for pathogenic variants on the other allele in the region of the reported deletion at chromosome 9q34.11, including the already disease-related genes (*DOLK*, *LRRC8A*, *NUP188* and *SET*). We had already identified that the 5' well conserved N-terminal

domain was present in the large *de novo* deletion (Figure 7.5, A; Appendix H, Figure H3). Additionally, our WES analysis identified a missense variant [NM_015354.2:c.3922C>T; NP_056169.1:p.(Arg1308Cys)] in *NUP188* present on the other allele affecting a well-conserved residue (Figure 7.5, B; Appendix H, Figure H4). Analysis of WES data for all other genes within the deleted did not reveal any other potentially pathogenic variants. Sanger sequencing confirmed the missense variant was inherited from the father (Figure 7.5; C). Although the identified heterozygous missense variant in *NUP188* is present in 13 individuals in gnomAD [total: 282468; mean allele frequency (MAF): $4.6*10^{-5}$] but no homozygotes have been listed in the database. All the reported individuals are of African ethnicity (aged between 40-80 years) whilst the patient was not of African ethnicity. Using *in-silico* tools we predicted this variant to be likely disease causing, due to high CADD score (31) and high Grantham score (180). Moreover, the preliminary 3D molecular modelling predicted this missense variant to cause loss of hydrogen bonding, disturbing the correct protein folding when the wild-type residue (Arg) was replaced with the comparatively smaller and more hydrophobic variant amino acid (Cys).

Western blot analysis of whole cell extracts prepared from individual 2 and control fibroblasts showed non-specific bands when probed with two different anti-SET antibodies, but significantly reduced protein levels were found with the anti-NUP188 antibody (Figure 7.5; D and E). Previous functional studies for NUP variants in affected individuals with neurological disorders have reported perturbed nucleoporin structure, enlarged nuclear size and impaired mRNA export (Fichtman et al., 2019; Theerthagiri, Eisenhardt, Schwarz, & Antonin, 2010). Our preliminary immunofluorescence studies did not reveal any significant differences in the overall nucleoporin structure between cells from a healthy control and individual 2 (Figure 7.5; F and G). This was determined by the intensity of mab414 an antibody that immunostains the nuclear pore complex proteins embedded in the nuclear envelope (Theerthagiri et al., 2010). Our preliminary RNA-FISH studies to determine whether there were mRNA import/export defects did not reveal any significant differences in the mean intensity ratio of nuclear and cytoplasmic accumulation of poly(A)+ RNA in fibroblasts cells from individual 2 (Figure 7.5; H and I). However, the nuclear size in individual 2 fibroblasts, determined using ImageJ, was significantly larger when compared to control fibroblasts (Figure 7.5; H and J).

Individual 3: KCTD16 variant

Using singleton WES, a novel homozygous variant [chr5:g.143853327T>A; NM_020768.3: c.937T>A, p.(Ser313Thr)] was identified in the *KCTD16* gene affected a well-conserved residue in the H2 domain in this female with classic RTT (Figure 7.6; A). Sanger sequencing confirmed this variant to be present in a heterozygous state in both of the healthy unrelated parents who were of Ashkenazi Jewish heritage (Figure 7.6; B). Although this variant is present in gnomAD in a heterozygous state at a very low MAF ($2.4*10^{-4}$; 68/282274 cases), no homozygotes were listed in the database. Moreover, most of these individuals listed in gnomAD were from the Ashkenazi Jewish population. Various *in-silico* tools (MutationTaster and SIFT) predicted this variant to be disease causing. Our 3D molecular modelling predicted the variant residue would disrupt proper protein conformation and so would be likely to destabilize the protein structure.

We undertook electrophysiology studies of the homozygous *KCTD16* variant [c.937T>A; p.(Ser313Thr)] identified in individual 3 using mature *Xenopus laevis* oocytes. KCTD16 is an auxiliary subunit of the GABA-B receptor. During synaptic transmission, activated GABA-B receptors cause release of G-proteins that facilitates postsynaptic slow inhibition through G protein-activated GIRK and inhibits voltage-gated calcium channels to reduce presynaptic transmitter release (Zheng, Abreu, Levitz, & Kruse, 2019). In our studies, the cRNA mixture containing human *GABBR1B* (or GABA-B1), *GABBR2* (or GABA-B2), *GIRK1* (or GIRK1), *GIRK2* (or GIRK2) were co-injected along with the cRNA of either wild type or variant [c.937T>A, p.(Ser313Thr)] *KCTD16*. A commonly occurring population variant [p.(Asp160Asn)] was also included as an internal control, as this variant was predicted to not change KCTD16 function. Preliminary TEVC recording performed after four days of incubation at 17-18ºC revealed no significant difference in K^+ current parameters, including raw current (µA) (Figure 7.6; C) and peak current (Figure 7.6; D); consistent with previous studies characterizing the wild type KCTD16 in *Xenopus* oocyte (Schwenk et al., 2010). In addition, parameters for receptor activation kinetics including the 20-80% rise time of K^+ currents (Figure 7.6; E), 50% onset time (Figure 7.6; F) and the tau constant from the first exponential fit (Figure 7.6; G) were also similar between wild type and variant KCTD16. These parameters will be further evaluated in a larger *Xenopus* cohort size to further evaluate this variant for potential pathogenicity. In addition, an alternative Chinese Hamster Ovary (CHO)

cell-based system will be used. The advantage of using CHO cells is that the constructs will be expressed in a mammalian cell-line, and they accommodate easier transfection, do not express endogenous GABA receptors, have low background conductance and allow performance of assays in a large number of adhered cells (Gamper, Stockand, & Shapiro, 2005; Schwenk et al., 2010). Interestingly, previous studies have shown that GABA-B response towards KCTD16 was dependent upon agonist concentration (Schwenk et al., 2010). Therefore, in future experiments, we will analyse the dose-response relationship of the GABA-B receptor with an increasing concentration of baclofen, the GABA-B agonist, in both wild type and variant KCTD16 expressing CHO cells.

Table 7.2: Variant curation and classification

Category	Individual 1	Individual 2			Individual 3
Genetic screening	Singleton, WGS	Trio; Whole genome SNP microarray		Singleton, WES	Singleton, WES
Gene symbol	*CHD8*	*SET*	*NUP188*	*NUP188*	*KCTD16*
Variant	chr14:21866016G>A; c.5017C>T; p.(Arg1673*)	9q34.11 (chr9:131,455,942-131,743,585): 3' region of *SET*	9q34.11 (chr9:131,455,942-131,743,585): 5' region of *NUP188*	chr9:131763886C>T; c.3922C>T; p.(Arg1308Cys)	chr5:g.143853327T>A; c.937T>A; p.(Ser313Thr)
Variant type	Non-sense	Deletion	Deletion	Missense	Missense
Zygosity	Heterozygous	Deletion	Heterozygous	Heterozygous	Homozygous
Segregation	Absent in mother	*de novo*	*de novo*	Inherited from father	Inherited from parents
Novel or reported	Novel	Novel	Novel	Novel	Novel
Affected /total exons	Exon 26/37	Exon (6-8)/8	Exon (1-14)/44	Exon 35	Exon 4/4
Stop codon (variant/ WT)	1673/2582	178/290	505/1749	1749/1749	428/428
Conservation	High	High	High	High	High
Affected protein domain	complete loss; Brahma and Kismet (BRK)	partial loss; nucleosome assembly protein (NAP)	partial loss; binding to Pom34	single amino acid change; unknown	H2 domain
Function of domain	Binds helicases	assembly of nucleosomes	Pom34p links NPC sub-complexes	Unknown	Unknown[$]
Population database (gnomAD)	Absent	Absent	Absent	4.6e-5; all African ethnicity	Absent
MutationTaster[#]	disease causing (p=1); NMD	disease causing (p=1); NMD	NA	disease causing (p=0.999)	disease causing (p=0.999)
PolyPhen-2[##]	NA	NA	NA	probably damaging (p=0.9120)	probably damaging (p=0.9560)
SIFT[###]	NA	NA	NA	damaging (0.00)	damaging (0.00)
CADD Phred-like score^	NA	NA	NA	31	23.8
Grantham score^^	NA	NA	NA	180	58
pLI constraint^^^/ Z missense score	1 (LoF)	1 (LoF)	0.95 (LoF)	1.12 (missense)	1.23 (missense)
Molecular modelling	Loss; Brahma and Kismet	NA	Loss; Binding to Pom34	Effect on protein folding	Effect on protein folding
ClinVar	Absent	Absent	Absent	Absent	Absent

. #: v2.0. URL = www.mutationtaster.org/; ##: v2.2.2r398. URL = http://genetics.bwh.harvard.edu/pph2/; Score range: 0.0 (tolerated) to 1.0 (deleterious); ###: v6.2.1. URL = http://sift.jcvi.org/; Score range: 0.0 (deleterious) to 1.0 (benign). Prediction cut-off = 0.05; ^CADD score: 0 (least damaging) to 100 (most damaging); ^^Grantham score: range 0 to 215: ≥100 = significant amino acid substitution; ^^^pLI scores : ≥ 0.9 (extremely LoF intolerant); ≤ 0.1 (LoF tolerant). $: reported to inhibit receptor desensitization (Fritzius et al., 2017; Seddik et al., 2012). NA = not available.

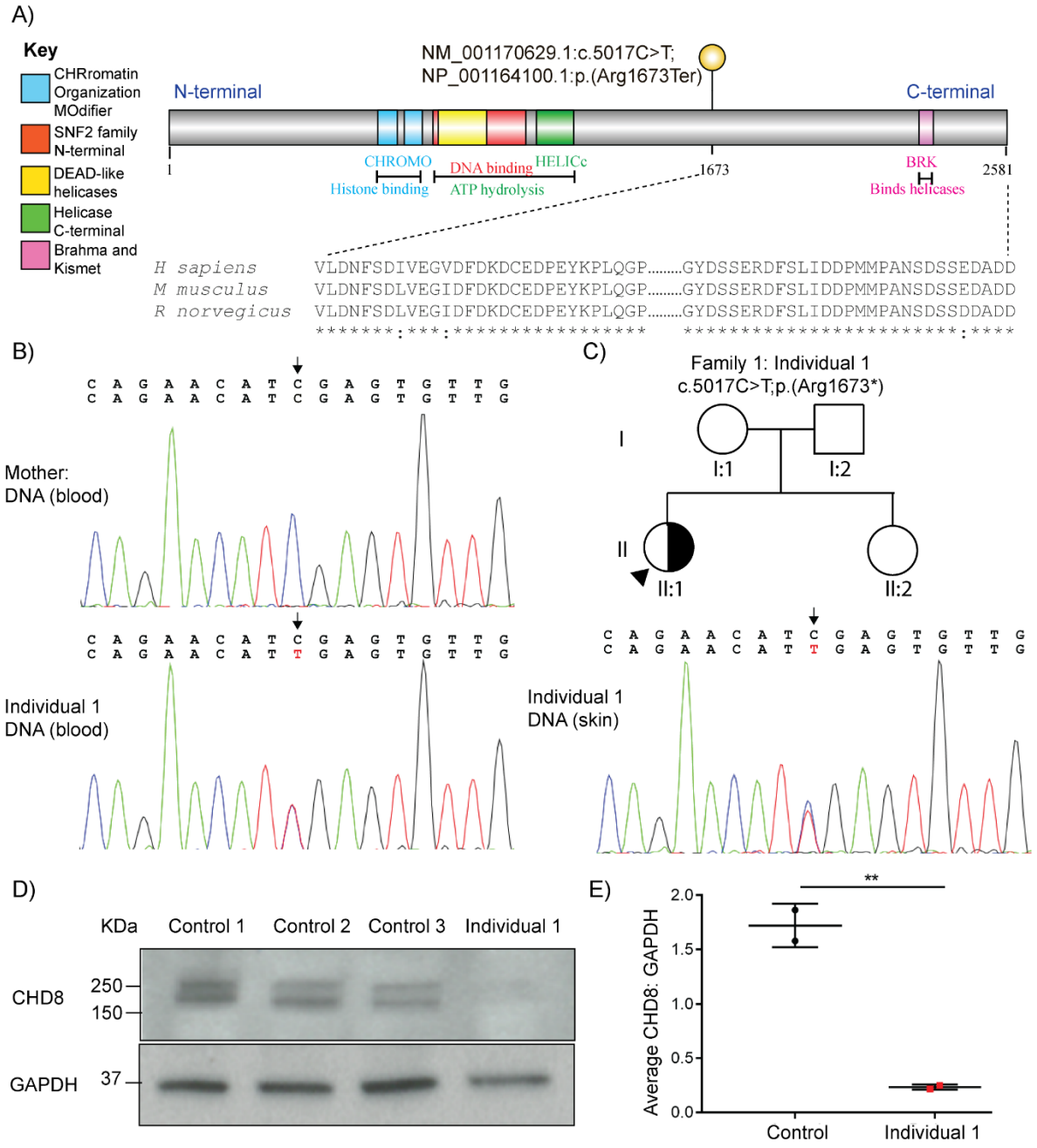


Figure 7.3: Schematic representing CHD8 domain structure, evolutionary conservation of C-terminal deletion region, Sanger confirmation and immunoblotting.

A) Protein domain structure of CHD8 (*H. sapiens*; NP_001164100.1) containing Chromatin organization modifier domain (CHROMO; 644-704 aa and 725-781 aa; cyan) involved in histone binding, SNF2 family N-terminal domain (SNF2_N; 814-1101 aa; red) crucial for DNA binding, DEAD-like helicases superfamily domain (DEXDc; 1133-1247 aa; yellow), Helicase conserved C-terminal domain (Helicase_C; 1133-1247 aa; green) responsible for ATP hydrolysis and Brahma & Kismet domain (BRK; 2310-2354aa; Pink) critical for binding

helicases. The novel heterozygous non-sense variant [NM_001170629.1:c.5017C>T; NP_001164100.1:p.(Arg1673*)] in individual 1; highlighted as yellow circle on CHD8 domain structure, results in the truncation of region between 1673-2581 amino acid. **B)** Multiple sequence alignment shows the sequence conservation at the start and end of the deleted C-terminal region of CHD8. Full alignment is available in Appendix H (Figure H1). The sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_001164100.1), *M musculus* (NP_963999.2), *R norvegicus* (NP_001334590.1). * (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a .(period) indicates conservation between residues with weakly similar properties. **C)** Family 1 pedigree and Sanger sequencing traces for mother and individual 1 with CHD8 variants. Circles = females and squares = males. Half-darkened symbol represents heterozygous change, arrow represents the affected individual. Sanger sequencing traces shown alongside in the pedigree shows that all the identified CHD8 variant is present in patient 1 DNA extracted from blood and primary skin fibroblast. However, the variant is absent in mother's DNA extracted from blood. Father DNA is unavailable to confirm segregation. **D) and E)** Western blot showing markedly reduced protein expression in individual 1 as compare to three independent control using anti-CHD8 antibody (Cat no.11891, Cell Signaling, 1:2000) targeting C-terminal domain of CHD8 downstream to the position affected by patient 1 variant. Experiment repeated two times with independent whole cell lysate preps. Data is represented as Mean±SD. Statistical test used was Man Whitney U test and ** represents $p<0.01$.

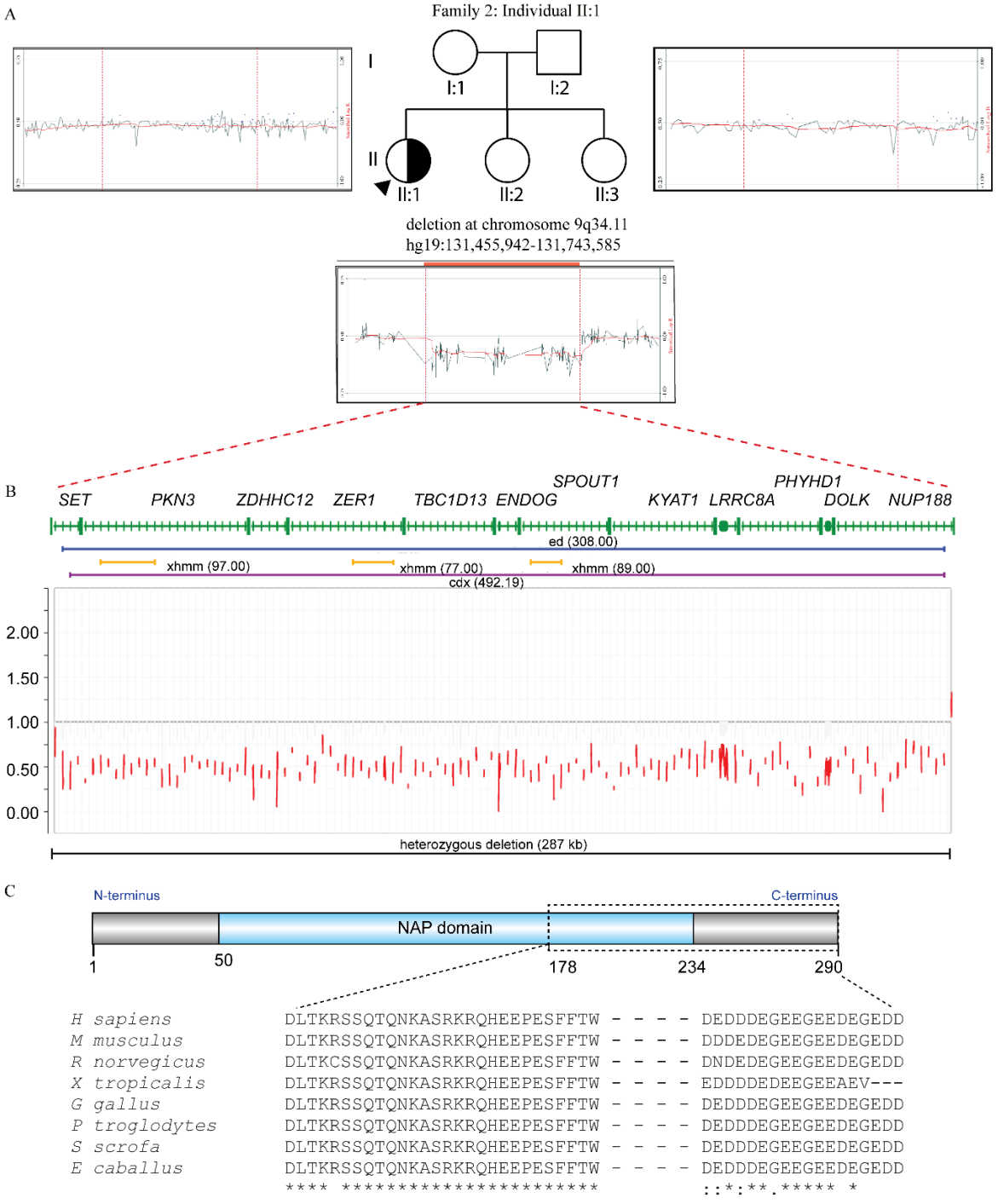


Figure 7.4: *De novo* deletion in individual 2, domain structure of SET and its protein levels expression in fibroblasts from individual 2 fibroblasts.

A) Pedigree of individual 2's family along with the molecular karyotyping showing the *de novo* deletion at chromosome 9q34.11 (hg19:131,455,942-131,743,585). More detailed information available in Appendix H (Figure H2). **B)** Schematic showing 13 genes including the 3' end of *SET* affected due to the large deletion. This deletion is independently confirmed through our singleton WES data as well using Ximmer tool (Sadedin, Ellis,

Masters, & Oshlack, 2018) based on different CNV detection algorithms: ed = exomeDepth (Plagnol et al., 2012), xhmm = eXome-Hidden Markov Model (Fromer et al., 2012), cdx = CODEX (Jiang, Oldridge, Diskin, & Zhang, 2015) **C)** Protein domain structure of SET (*H. sapiens*; NP_001116293.1) containing Nucleosome assembly protein domain (NAP: 50-234 amino acids; blue). The 3' region of SET truncated due to large deletion (178-290 amino acids) is highlighted using dotted rectangle. Multiple sequence alignment depicting the sequence conservation at the start and end of the deleted C-terminal region of SET. For full alignment, refer to Appendix H (Figure H3). The sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_001116293.1), *M musculus* (NP_076360.1), *R norvegicus* (NP_001012522.1), X tropicalis (NP_989041.1), *G gallus* (NP_001025862.2), *P troglodytes* (XP_024201677.1), *S scrofa* (NP_001231019.1), *E caballus* (XP_023484824.1). * (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a .(period) indicates conservation between residues with weakly similar properties.

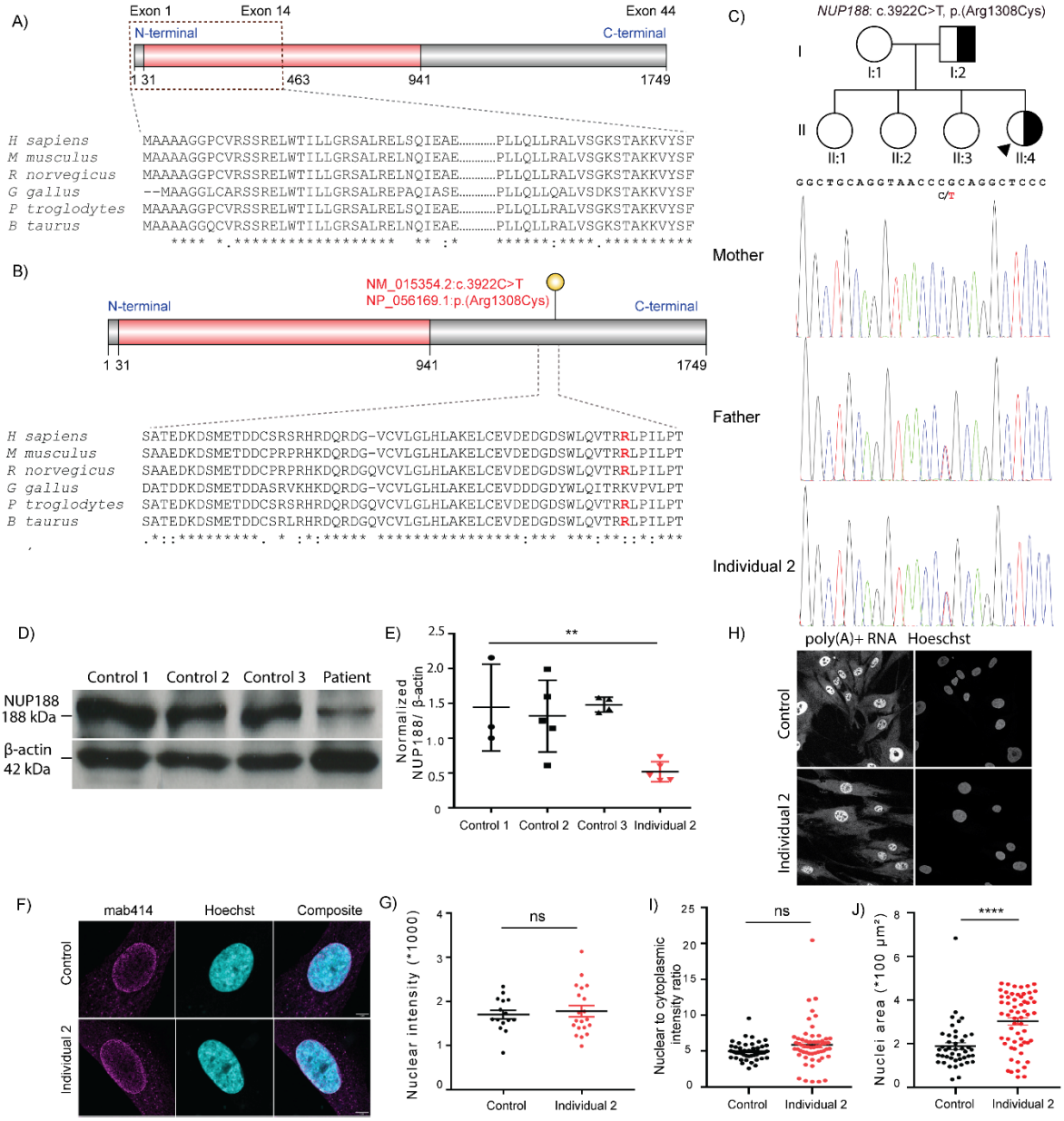


Figure 7.5: Compound heterozygous *NUP188* variants in individual 2, genetic analysis and preliminary functional validation findings.

Multiple sequence alignment shows the sequence conservation **A)** at the start and end of the deleted N-terminal region of NUP188 and **B)** in the region flanking the missense variant. The variant residue is highlighted in red. The sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_056169.1), *M musculus* (NP_938046.2), *R norvegicus* (XP_006233967.1), *G gallus* (NP_001025861.1), *P troglodytes* (XP_016801867.2), *B taurus* (NP_001095308.1). * (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a .(period) indicates conservation between residues with weakly similar properties. **C)** Family 2 pedigree and Sanger sequencing traces for *NUP188* missense variant. Circles = females and squares = males. Half-darkened symbol represents heterozygous change, arrow represents the affected individual. Sanger sequencing traces shown below

pedigree showing identified *NUP188* variant in individual 2 being inherited from father. **D and E)** Western blot showing significantly reduced protein expression in individual 2 as compared to three independent control using anti-NUP188 antibody (Cat no. A302-322A-M, Bethyl Lab, 1:1000)–targeting C-terminal domain of NUP188 (region between residue 1600 and 1650). Experiment was repeated using three independent batches of whole cell lysates prepared from controls and individual 2 fibroblasts. Data was analysed using ordinary one way ANOVA. **F and G)** Preliminary findings of the immunofluorescence performed on skin fibroblasts using single control and individual 2 using mab414 antibody (Cat. no. ab24609, Abcam; 1:500) and Hoechst stain to label nuclei. The monoclonal mab414 antibody revealed well-defined punctuate staining on the nuclear rim. The fluorescence intensity measured along the perimeter of the nucleus in 20 cells did not show any significant difference. **H, I and J)** Preliminary RNA-FISH performed on one control and individual 2 fibroblasts. No difference in the intensity ratio between nucleus and cytoplasm was observed. However, a significantly increased nuclei area was observed in case of individual 2's fibroblasts (n= 60 cells) as compared with the controls (n=44 cells). Data represented as mean ± SD was analysed using Mann-Whitney U test. ns = not significant; ** represents $p \leq 0.01$; **** represents $p \leq 0.0001$.

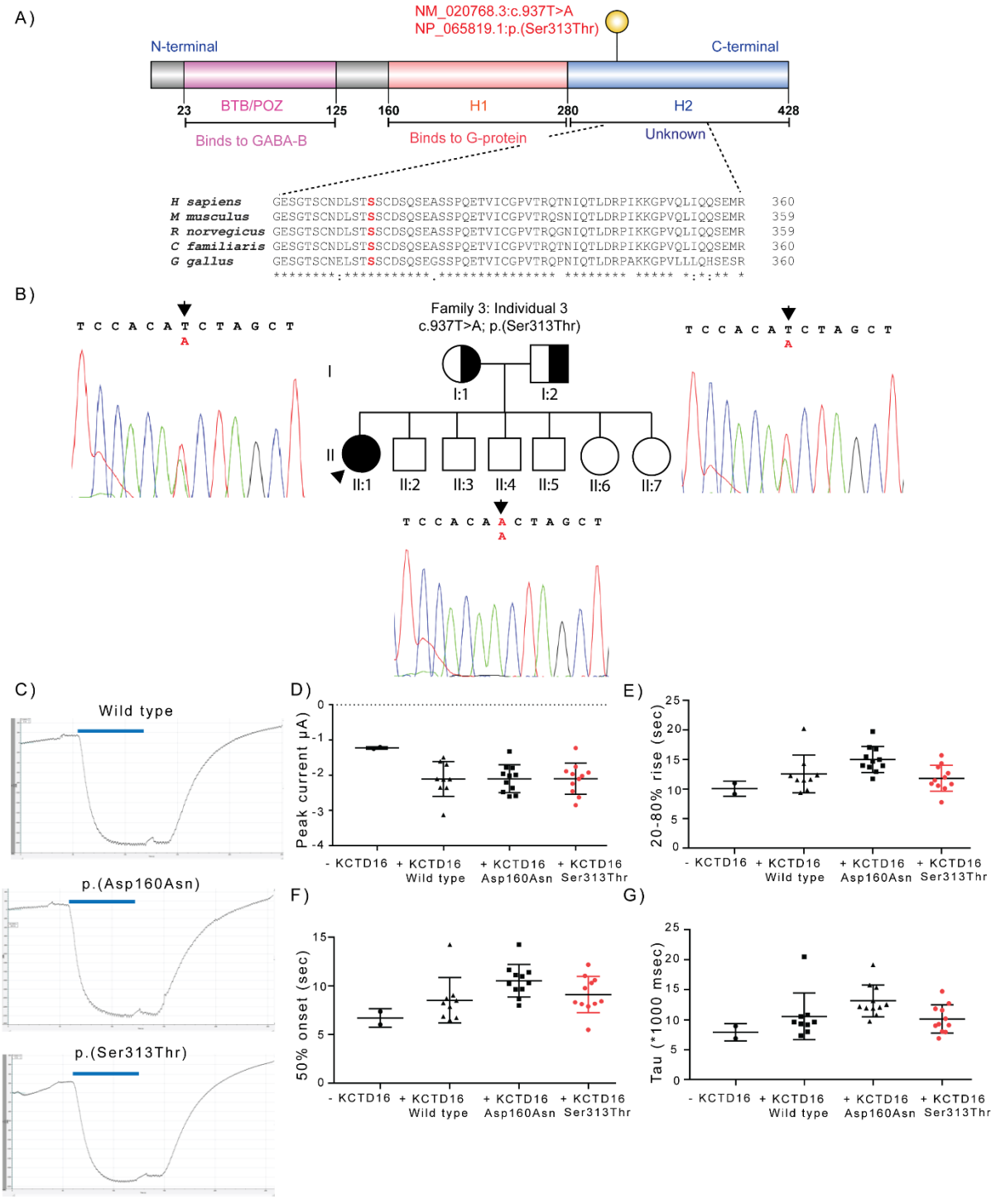


Figure 7.6: KCTD16 domain structure, location of variant and preliminary findings from functional validation using *Xenopus* oocytes.

A) The domain structure of KCTD16 consisting of two main domains: BTB/POZ domain (amino acid 23-125; pink) that binds to the GABA-B receptor and H1 domain (amino acid 160-280; red) that associated with the G-protein. The homozygous variant [chr5:g.143853327T>A; NM_020768.3:c.937T>A; NP_065819.1:p.(Ser313Thr)] identified in the RTT individual 3 was found to be localized in the domain with

unknown function (H2; amino acids 280-428; blue).B) Sanger sequencing confirmed that both parents are heterozygous for the change which is identified as homozygous in affected individual 3, thus confirming the segregation of the *KCTD16* variant **C)** Representative raw current traces from each expression group. Blue line indicates baclofen (1 μM) application. GABA receptor response was evaluated in terms of K^+peak current (μA) **(D)**, 20-80% rise time (sec) **(E)**, 50% onset (sec) and time weighted constant tau **(msec)** (M. Li et al., 2017; Schwenk et al., 2010). No significant differences were identified.

7.4 Discussion

A significant number of individuals with NDDs including RTT do not necessarily have a confirmed genetic diagnosis. This limits the complete understanding of underlying biology of the disorder in those individuals, preventing them from gaining access to targeted therapies. Impaired brain development due to altered regulatory expression of specific genes has been reported as a common feature among RTT/ RTT-like individuals (Kumar et al., 2019; Tyagi et al., 2016). A number of regulatory genes implicated in these pathways has been now recognized to cause RTT. It is noteworthy that NGS is continuously uncovering genetic links between RTT and several other molecular pathways. Thus, molecular characterization of key proteins involved in these pathways is critical to understand the complete patho-physiology behind these disorders.

7.4.1 CHD8

As stated above, perturbed gene expression regulation is a common pathway underlying NDDs including RTT (Kumar et al., 2019; Tyagi et al., 2016). Pathogenic variants in several proteins that alter histone modifications, remodulate the chromatin structure and change the transcriptional active states of various genes have been implicated in NDDs (Iwase & Martin, 2018). The nine members of the ATP-dependent chromatin remodelling complex, Chromodomain Helicase DNA binding protein (CHD) family, classified into three subgroups, are of key interest during early neurodevelopment processes owing to their broad regulatory role in neuronal migration and differentiation (Nitarska et al., 2016; Sokpor, Castro-Hernandez, Rosenbusch, Staiger, & Tuoc, 2018). Even though CHD7 and CHD8 belong to same subgroup III of the CHD family of proteins, they play distinct roles in nucleosome binding and remodelling (Manning & Yusufzai, 2017). Pathogenic variants in *CHD7* cause an autosomal

dominant disorder CHARGE syndrome, characterised by growth/development retardation and a specific pattern of congenital defects including choanal atresia and abnormalities of the heart, inner ear, and retina (Hsu et al., 2014). Truncating variants in *CHD8* have been reported to be one of the strongest risk factors for autism spectrum disorders (ASDs), a heterogeneous group of disorders with a complex genetic eitology (Ayhan & Konopka, 2019; Bernier et al., 2014; Stolerman, Smith, Chaubey, & Jones, 2016). Individuals with *de novo* heterozygous *CHD8* pathogenic variants have intellectual disability, developmental delay, macrocephaly, facial dysmorphisms, motor delay, hypotonia, gastrointestinal problems and abnormal cortical development (Cotney et al., 2015). Moreover, homozygous deletion of *Chd8* in mice was embryonic lethal (Nishiyama et al., 2009). However, heterozygous mice were displayed ASD-like behaviours such as abnormal social interaction, increased anxiety and reduced social memory, consistent with the phenotype often observed in patients with pathogenic variants in *CHD8* (Gompers et al., 2017; Katayama et al., 2016). *Chd8* haplo-insufficient mice also demonstrated abnormal expression of genes implicated in chromatin modification and cell cycle regulation (Gompers et al., 2017). Recently, the *Drosophila* homolog of CHD7 and CHD8, known as Kismet, was identified to be involved in synaptic vesicle recycling during endocytosis, and in regulation of stem cell proliferation (Gervais et al., 2019; Latcheva et al., 2019). Individual 1 in our study exhibited several clinical features that are often observed in individuals with *CHD8* defects, however she also fulfilled the criteria for atypical RTT (Neul et al., 2010). Although we could not confirm segregation of the non-sense variant in *CHD8* in the father, the reduction in CHD8 protein levels implied that the variant in *CHD8* may explain the clinical phenotype (Bernier et al., 2014).

7.4.2 Chromosome 9q34.11 deletion including *SET* and *NUP188* genes

Of the 13 genes present in the *de novo* deletion (hg19:131,455,942-131,743,585) in individual 2, only four genes (*DOLK*, *LRRC8A*, *NUP188* and *SET*) have been listed as disease related in the reported literature. Homozygous variants in Dolichol Kinase (*DOLK*; OMIM: 610746) result in the autosomal recessive congenital disorder of glycosylation type IM, characterised by muscular hypotonia, dilated cardiomyopathy and early death (Helander et al., 2013; Kranz et al., 2007; Lefeber et al., 2011; Lieu et al., 2013). Our WES data did not identify a second pathogenic variant in *DOLK* on the other allele, and as the clinical phenotype is distinct from

that of individual 2, we can say that the heterozygous deletion in *DOLK* is not contributing to the phenotype. Another gene present in the deleted region in individual 2, Leucine rich repeat containing 8 VRAC subunit A (*LRRC8A*; OMIM: 608360), is implicated in autosomal dominant agammaglobulinemia 5, characterised by a weakened immune system, reduced B-cell lymphocytes and dysmorphic facial features including high-arched palate, epicanthic folds, mild hypertelorism, and low set ears (Sawada et al., 2003). However, no recurrent infections or any other immune-deficiency disorder was reported in individual 2. The testing of gamma globulin levels has not been performed yet on her so it may be possible that she is at risk of immune problems.

The large deletion in individual 2 also included a known disease gene called SET nuclear proto-oncogene (*SET*; OMIM: 600960), also known as PP2A Inhibitor 2 (I2PP2A), that encodes a nuclear phosphoprotein (Lam et al., 2012). The possible role of SET in NDDs was first revealed when a heterozygous protein truncating variant was identified in a 12 year old male with global developmental delay, moderate intellectual disability, no autistic features, attention deficit without hyperactivity, severe speech delay, congenital microcephaly, dysmorphic facial features and a normal brain MRI (Hamdan et al., 2014). Additionally, trio WES identified individuals (two male and one female) exhibiting speech delay, normal head circumference and behavioral issues with likely pathogenic frameshift variants in *SET*. Interestingly, all patients had mild dysmorphic features including a wide mouth, thickened lower lip and vermillion, broad base of the nose and widely spaced teeth (Deciphering Developmental Disorders, 2017; Richardson et al., 2018). Moreover, through WES, Stevens et al. (2018) had identified likely pathogenic heterozygous missense, protein truncating or splicing variants in *SET* gene in five individuals affected with autosomal dominant non-syndromic intellectual disability with motor and speech delay. All the affected individuals carried a *de novo* heterozygous variants except for one in which maternal inheritance was noted (Stevens et al., 2018). Interestingly, similar clinical features were observed in affected individuals with protein truncating and missense variants in SET (Stevens et al., 2018) that overlapped with our RTT-like individual 2 (Appendix H; Table H3).

SET exhibits multiple chromatin regulatory functions including nucleosome assembly, transcription, histone chaperoning and gene transcription (Park & Luger, 2006) by inhibiting histone acetyltransferases (HATs) dependent histone H4 acetylation (Cervoni et al., 2002; Wang et al., 2016). SET interacts with several other histone modification proteins that have been implicated in autosomal dominant intellectual disability including SET-binding protein 1 (SETBP1), E1A-binding protein, 300kDa (EP300), Lysine Methyltransferase 2A (KMT2A), cAMP-response element binding protein (CREBBP), Ras-related C3 botulinum toxin substrate 1 (RAC1) and CTCF1 (S. J. C. Stevens et al., 2018). In addition, SET plays a critical role in neurogenesis by regulating neurite outgrowth, (Lam et al., 2013; Klooster et al., 2007; Wang et al., 2019) and by binding with the neuron-specific cyclin dependent kinase 5 activator, critical for neuronal differentiation (Nikolic et a1. 1996; Ohshima et al., 1996; Qu et al., 2002), and GABA/ N-methyl-D-aspartate receptor regulation (Compagnone et al., 2000). SET is highly expressed in neural tubes and *SET* homozygous null mouse were embryonic lethal, with reduced embryo size, cardiac edema and open neural tube defects (Compagnone et al., 2000). Together, these studies indicate the critical role of SET in the transcriptional regulation of genes critical for neurogenesis and neuronal differentiation.

Functional haplo-insufficiency in SET could be due to loss of the highly conserved NAP domain, critical for chromatin regulation (Park & Luger, 2006). The *de novo* deletion identified in individual 2 resulted in loss of the last three exons (exon 6, 7 and 8), predicted to cause partial truncation of the NAP domain (Lindeboom et al., 2016), Despite attempts at western analysis of control and individual 2 fibroblasts with two different SET antibodies, only non-specific bands were detected. Future studies will perform westerns using other SET antibodies. However, given the previous reports of pathogenic *SET* variants, we summarize that the protein truncating variant affecting NAP domain in SET is likely to be contributing to the clinical phenotype of individual 2.

Of further interest in individual 2 was the missense variant (c.3922C>T; p.(Arg1308Cys) in one allele, and the partial truncation of NUP188 from the large deletion in the other allele.

Variations in NUPs are slowly emerging to be associated with disorders overlapping with NDDs. Biallelic missense and/or frameshift pathogenic variants in Nucleoporin, 214-kDa (*NUP214*; MIM: 114350) are associated with neurodevelopmental regression, seizures, myoclonic jerks, progressive microcephaly, and cerebellar atrophy (Fichtman et al., 2019). A heterozygous spile-site variant identified in an individual with autosomal dominant cardiovascular disease (Haskell et al., 2017) and knockdown of nup188 in *Xenopus* resulting in cardiac defects (Fakhro et al., 2011), suggests the role of NUP188 in heart development and function. In addition, homozygous protein truncating variants in NUP188 resulted in NDDs characterized with microcephaly, trigonocephaly, heart defects, congenital bilateral cataract, microphthalmia and early death (Sandestig et al., 2019). Brain MRI in previously reported individuals showed loss of periventricular white matter, thin corpus callosum and delayed myelination, with dysmorphic facial features including laterally extended arched eyebrows, a wide convex nose, a wide prominent nasal bridge and prominent angulated antihelices of the ears (Sandestig et al., 2019).

The deletion in individual 2 at chromosome 9q34.11 resulted in the loss of 5' end of the *NUP188,* corresponding to a region between exon 1 and exon 14, containing that resulted in lack of NUP188 protein expression. The N-terminal domain of NUP188 is involved in binding to the C-terminus of Pom34p, which is the transmembrane subunit of nucleoporin complex (NPC), and links NPC sub-complexes together (Miao et al., 2006; Theerthagiri et al., 2010). NPC is a large multi-protein octagonal multilayered channel (~50nm) across the nuclear envelope, consisting of several copies (~600) of evolutionary conserved Nucleoporins (NUPs) family consisting

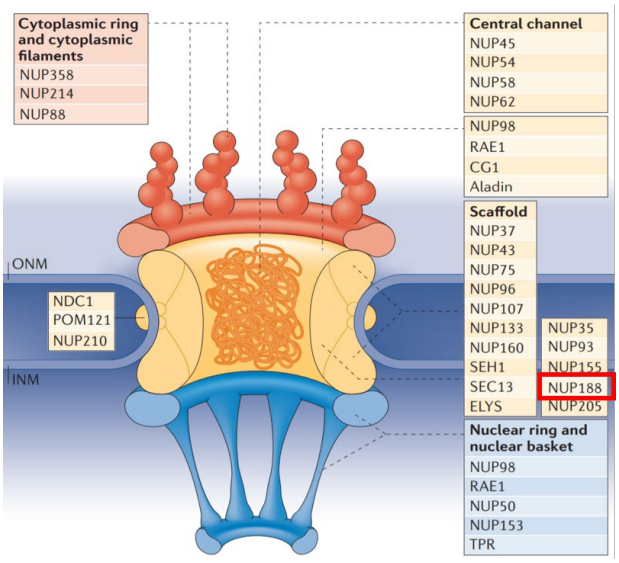


Figure 7.7: Nuclear pore complex (NPC) structure.

Schematic showing the key NUPs forming the structure of NPC including cytoplasmic ring, central channel, nuclear ring and scaffold. NUP188 (in red box) is an essential subunit in the scaffold that is critical to bind NPC with the nuclear membranes. Image from (Raices & D'Angelo, 2012).

of 30 protein members (Figure 7.7) (Theerthagiri et al., 2010). NPC acts as a "gatekeeper" that controls the nucleocytoplasmic active transport of large macromolecules as well as passive flow of small molecules. In addition to regulation of nuclear transport, NPCs are associated with chromatin organisation and gene regulation (D'Angelo et al., 2012; Liu et al., 2009; Yasuhara et al., 2007).

Previous functional studies of NUP214, a member of the NUP family, showed a normal density of NPC, but reduced protein levels and mRNA/ protein transport defects across the nuclear membrane, resulting in blocked NPCs, referred to as a "central plug" phenotype, in affected individual fibroblasts (Fichtman et al., 2019). Our preliminary functional studies showed that individual 2's fibroblasts cells contained structurally normal NPC's as evident from typical mab414 staining, however reduced NUP188 protein levels along with a comparatively increased size of the nucleus was observed. This is consistent with previous studies (Theerthagiri et al., 2010) that suggest that impairment in NUP188 function could lead to perturbed nuclear transport (Theerthagiri et al., 2010). Our preliminary RNA-FISH results did not show significant changes in total mRNA accumulation in the nucleus versus cytoplasm in individual 2 compared with controls. Future experiments will be repeated with an increased number of cells for comparative analysis to examine for more subtle effects. We also plan to investigate NPC density, previously shown to be increased as a result of loss of Nup188 (Theerthagiri et al., 2010), and to test for transport defects of key proteins involved in NDDs including RTT. In addition, we are currently developing tools to detect changes in the width and "central plug" phenotype of single NPC complexes. Thus, based on our initial clinical and functional evaluation of individual 2, it is plausible that changes in SET and NUP188 are both contributing to the clinical phenotype observed in individual 2.

Finally, in one classic RTT individual, we have identified a novel homozygous missense variant [c.937T>A; p.(Ser313Thr)] in *KCTD16*, for which *in silico* tools predicted pathogenicity. KCTD16 belongs to a large superfamily consisting of 26 members (Z. Liu, Xiang, & Sun, 2013). Although the role of KCTD16 in human disease is yet to be identified, other family member proteins, including KCTD3, KCTD7 and KCTD13 have been implicated in NDDs including intellectual disability, ASDs and epileptic encephalopathies (Teng et al.,

2019). Moreover, the GABAergic signalling pathway is one of the key pathways commonly perturbed in a number of NDDs including ASDs, Angelman syndrome and RTT (Cellot & Cherubini, 2014; Chattopadhyaya & Cristo, 2012; Di Cristo et al., 2011; Sgado et al., 2011). Studies have revealed that *Mecp2*-knockout mice exhibit functional defects in GABAergic signalling, particularly in the formation and maturation of GABAergic connectivity in neurons (Chao et al., 2010; Chao et al., 2007; Dani et al., 2005; Dani & Nelson, 2009). Interestingly, only a subtle reduction (~30-40%) in the GABA neurotransmitter release resulted in severe neurological phenotypes including stereotypic behaviours, gait defects, learning/ memory dysfunction that are commonly observed in RTT individuals (Chao et al., 2010).

7.4.3 KCTD16

KCTD16 is one of the KCTD family of proteins (Group D members: KCTD8, KCTD12, KCTD12b and KCTD16) and form the auxiliary subunit of GABA-B receptors (Figure 7.8). KCTD16 is exclusively expressed in the brain and regulates the kinetics of the receptor response for neuronal communication (Fritzius & Bettler, 2019). All KCTD family proteins contain a well-conserved BTB domain, similar to tetramerization T1 domain, that associate with the GABA-B2 receptor (Seddik et al., 2012). While KCTD12 contains a conserved homology H1 domain for receptor desensitization, KCTD8 and KCTD16 have an additional well-conserved H2 domain that inhibits receptor-desensitization (Seddik et al., 2012). While expression of KCTD16 resulted in subtle changes in the K^+ currents

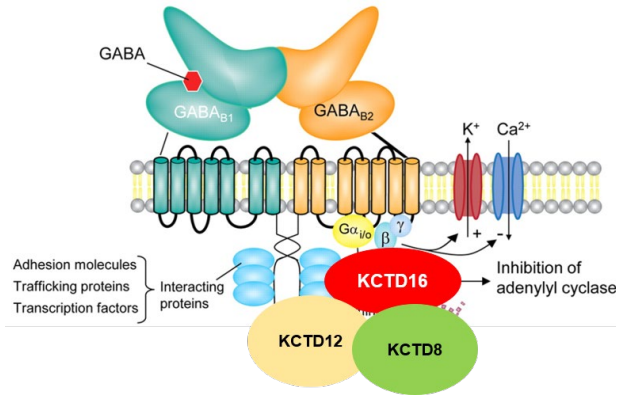


Figure 7.8: GABA receptor structure associated with KCTD proteins

Schematic showing the heterodimer GABA-B receptors formed with GABA-B1 (green) and GABA-B2 (orange) subunits. Auxiliary KCTD proteins, KCTD12 (yellow), KCTD8 (light green) and KCTD16 (red) associate with cytoplasmic tail of the GABA-B2 subunit to regulate K^+ and Ca^{2+} channels.

in *Xenopus* and CHO cells, KCTD12 caused a marked receptor desensitization response (Schwenk et al., 2010).

The characterization and functional evaluation of variants affecting KCTD family genes are routinely performed in alternative cellular models. In *Xenopus* oocytes and cultured hippocampal neurons, significant receptor desensitization was only observed for KCTD12, while KCTD16 showed only subtle changes. However, the GABA-B receptor activated 20–80% rise time of K^+ currents through GIRK1/ GIRK2 channels upon baclofen (GABA agonist) application in CHO cells revealed a marked increase in receptor response in all KCTDs. Interestingly, the extent of the response was dependent upon the subtype of KCTD (Schwenk et al., 2010). Moreover, KCTD12 and KCTD16 were found to lower the baclofen induced EC_{50} value (concentration required for half-maximum response of GABA-B receptor), suggesting that KCTD proteins may affect the kinetics and agonist potency of the GABA-B receptor (Schwenk et al., 2010). Our preliminary functional studies to evaluate the pathogenicity of the *KCTD16* variant in individual 3 with classic RTT in *Xenopus* oocytes showed no significant differences for receptor activation (K^+ current), receptor desensitization (20-80% rise time), rate of receptor response (50% onset) and decay response (tau) of GABA-B receptor for the *KCTD16* variant. It may be that the *KCTD16* variant is causing a subtle change in the GABA receptor kinetics that is not being detected in the currently studied *Xenopus* model. Thus, in line with previously published studies, we aim to undertake these studies in an alternative mammalian CHO cell-based model to ascertain the functional significance of the *KCTD16* variant.

In summary, this chapter has described the findings related to four additional identified genes (*CHD8*, *SET*, *NUP188* and *KCTD16*) that may contribute to the phenotypes observed in these RTT and RTT-like individuals. Planned experiments will further validate the pathogenicity of these variants prior to publication. Contingent on our ongoing work, we propose that these genes should also be considered as potential disease -causing genes in *MECP2* negative RTT/ RTT-like individuals.

Chapter 8

Chapter 8 Conclusions and Future Directions

8.1 Conclusions

RTT is a devastating neurological disorder (NDD) that can be clinically classified into two main categories; classic and atypical; based upon the published revised diagnostic criteria for RTT syndrome (Neul et al., 2010). Although, RTT is mainly caused by pathogenic variations in *MECP2*, between 3 – 14% of classic and atypical RTT individuals still do not have a genetic diagnosis (Neul et al., 2014). A small proportion of RTT individuals have also been found to carry pathogenic variants in other genes such as cyclin-dependent kinase-like 5 (*CDKL5*) and Forkhead box protein G1 (*FOXG1*), however pathogenic variants in these genes are now widely recognized to cause clinically distinct disorders from RTT (Hector et al., 2017; Kortum et al., 2011; Mitter et al., 2018). The significant clinical overlap between RTT and other NDDs further challenges the clinical diagnosis of mutation-negative affected RTT individuals, prompting a need to identify the genetic cause for their condition. Over the course of the past few years, many affected individuals with RTT features and additional features overlapping with other distinct NDDs, have been identified and are increasingly referred to as RTT-like individuals (Ehrhart et al., 2018; Iwama et al., 2019). Advanced and affordable NGS techniques have been gradually adopted to screen RTT/ RTT-like individuals in order to establish their genetic diagnosis. As a result, an increasing number of genes previously implicated in other NDDs are now being recognized to cause RTT/RTT-like features (Arafat et al., 2017; Jang et al., 2015; Lee et al., 2016; Lopes et al., 2016; Saitsu et al., 2014; Sajan et al., 2017). Despite this, there are still many RTT/RTT-like individuals with no genetic diagnosis. We aimed to collect DNA samples from such mutation-negative RTT/ RTT-like individuals in order to interrogate their genomes for likely disease-related genes.

In order to uncover the causative genes in genetically undiagnosed individuals with RTT/ RTT-like features, DNA samples were collected from 88 affected individuals through global collaborations. Based upon the resources available at the time of the study, we screened 42 individuals using a range of NGS techniques including WGS, WES or a targeted sequencing panel that contained a customised list of 136 genes associated with RTT and other related NDDs.

Sanger sequencing was then used to establish the segregation of the identified variants. Of the 42 affected individuals, two affected females (2/42; ~7%) were screened using WGS, eight by WES (8/42; ~19%) and 32 (32/42; ~74%) using our custom designed targeted sequencing panel. While WGS and WES identified potentially causative genes in almost all the individuals analysed (WGS=2/2; 100% and WES=7/9; 88%), the targeted sequencing panel approach had a much lower success rate (6/32; ~19%). However, in some cases, the panel approach may be more successful than WES and WGS due to the higher depth of sequencing of a focused list of genes carefully selected based on previous disease associations. An example of the successful utilisation of the custom gene panel was in the successful genetic diagnosis of a mosaic variant in *MECP2* that we identified in a classic RTT male that was previously missed using WES sequencing.

Overall, this research project has explored additional possible cellular pathways including microtubule associated pathways, chromatin remodelling, GABA signalling and protein translation pathways that may be perturbed in RTT/ RTT-like individuals. Molecular characterization of these pathways is critical to more completely understand the complex nature of RTT. This knowledge may also be extrapolated to better understand the biology behind other complex NDDs.

8.1.1 Microtubule associated pathways

Over the past few years, accumulating evidence implicated microtubule instability, and associated impaired neuronal trafficking of cargo by molecular motor proteins as potentially underlying the pathogenesis of NDDs, including RTT (Baltussen et al., 2018; Barbiero et al., 2017; Chahrour et al., 2008; Delepine et al., 2016; Delepine et al., 2013; Gold et al., 2015; Jordan et al., 2007; Kaufmann et al., 1995; Maezawa et al., 2009; Nectoux et al., 2012; Williams et al., 2014; Xu et al., 2014). Through this project, we have provided functional evidence that pathogenic variants in *KIF1A* lead to defective KIF1A mobility that may contribute to RTT pathogenesis (Chapter 3). KIF1A is a neuron-specific kinesin motor protein critical for anterograde axonal transport of synaptic cargo. Variations in *KIF1A* are associated with a spectrum of neurological conditions, collectively known as KAND (Boyle &

Chung, 2017). Our research has expanded the phenotypic spectrum of KAND to include RTT/ RTT-like clinical phenotypes. We suggest that the screening of *KIF1A* should be undertaken in genetically undiagnosed RTT/RTT-like individuals. Furthermore, we have evaluated the functional consequences of four variants affecting the motor domain of KIF1A in our RTT cohort (Chapter 3). We have also characterised the functional consequences of 16 *KIF1A* variants in 27 KAND individuals exhibiting complex and variable phenotypes. These variants had not been previously analysed at a functional level (Chapter 4). Our study provides further evidence that although different variants in *KIF1A* may result in a phenotypic overlap, specific clinical features and the clinical severity may be influenced by other factors such as the location of the variant. Along with expanding the genotype-phenotype correlation associated with *KIF1A* variants, this study has opened up avenues for drug discovery approaches to improve the function of KIF1A in individuals affected with RTT or KAND.

8.1.2 Chromatin remodelling pathways

Recent NGS studies for RTT have uncovered a large number of genes involved in a range of biological processes, and a significant proportion of these genes belong to the epigenetic regulatory pathways (Jang et al., 2015; Sajan et al., 2017). Impaired chromatin regulation is a common molecular mechanism established for several NDDs including RTT (Kumar et al., 2019; Tyagi et al., 2016). Moreover, the evidence of common pathways between RTT associated genes (*MECP2*, *CDKL5* and *FOXG1*) involved in regulating chromatin structure further suggests an underlying association between epigenetic regulation defects and RTT (Carouge et al., 2010; Mari et al., 2005). Our study has further highlighted the possible involvement of an additional four chromatin regulator genes (*KAT6A*, *SMC1A*, *CHD8* and *SET*) in RTT/ RTT-like individuals (Chapter 5 and Chapter 6). Of these four genes, variants in *SMC1A*, often implicated in CdLS, have previously been associated with several RTT/RTT-like individuals (Gilissen et al., 2014; Huisman et al., 2017; Sajan et al., 2017). Our study has further expanded the genetic landscape of RTT to include three new chromatin regulator genes (*KAT6A*, *CHD8* and *SET*) that have been previously associated with other NDDs, but were not recognized to contribute to RTT/RTT-like phenotypes.

8.1.3 GABA signalling pathway

Perturbed GABA signalling, previously implicated in various NDDs including Angelman syndrome and ASDs, has been also found to be defective in *Mecp2*-knockout mice, highlighting its contributing role in RTT (Chao et al., 2010; Chao et al., 2007; Dani et al., 2005; Dani & Nelson, 2009). In our study of RTT/ RTT-like individuals, potentially pathogenic variants were identified in two genes (*SHANK3* and *KCTD16*), both involved in GABA signalling and synaptic transmission (Chapter 5 and Chapter 6). *SHANK3* is one of the most common genes associated with ASDs and known to contribute to Phelan-McDermid syndrome, a rare NDD characterized by intellectual disability, developmental delay, abnormal muscle tone and autism related clinical features (De Rubeis et al., 2018). While *SHANK3* has been previously associated with RTT/ RTT-like syndrome (Hara et al., 2015), pathogenic variants in *KCTD16* have not been linked to any human disease. Interestingly, biallelic variations in other related members of the KCTD family (*KCTD3*, *KCTD7* and *KCTD13*) are now known to cause NDDs including intellectual disability, autism, seizures and a severe early onset progressive disorder with epilepsy (Teng et al., 2019). We suspect that the homozygous variant identified in our classic RTT proband is pathogenic and contributes to individual's phenotype. The preliminary electrophysiology studies using a *Xenopus* oocyte model did not identify any significant differences in the electrical activity properties between the wildtype and individual is specific variant in *KCTD16*. However, we aim to use more relevant mammalian Chinese hamster ovary (CHO) cell-based model with alternate G-protein signalling pathway that is not limited by the delay in solution exchange issues that are often seen in *Xenopus*.

8.1.4 Protein translation pathways

Regulation of protein translation in polarized neurons is critical for the local synthesis of several proteins essential for neurite outgrowth and to maintain synaptic plasticity, both of which are critical pathways in establishing functional neuronal networks in brain (Van Driesche & Martin, 2018). Thus, it is not surprising that defective protein translational control has been found to cause several NDDs including intellectual disability, ASDs, and epilepsy related conditions (Scheper, van der Knaap, & Proud, 2007; E. T. Wang et al., 2016). A key gene. *EEF1A2*, involved in protein translation has been related with autism, intellectual disability and early infantile epileptic encephalopathy (de Ligt et al., 2012; Epi K consortium

& Epilepsy Phenome/Genome Project, 2017; Inui et al., 2016; Iossifov et al., 2014; Lam et al., 2016; Lelieveld et al., 2016; Lopes et al., 2016; Nakajima et al., 2015; Ohba et al., 2015; Veeramah et al., 2013). So far, in the literature only a single RTT-like female has been reported to carry a likely pathogenic variant in *EEF1A2*. Through our study, we have identified an additional *EEF1A2* variant, previously linked with NDDs, in a RTT-like individual. Our findings further strengthens the casual association between impaired EEF1A2 function and RTT (Chapter 6).

8.1.5 Nucleoporin based pathways

Interestingly, we have identified in a RTT-like female a large *de novo* deletion encompassing 13 genes, of which two genes (*SET* and *NUP188*) are likely contributing to her phenotype (Chapter 6). Several pathogenic variants in the chromatin regulator gene *SET* have been previously identified in individuals with clinical features overlapping with our RTT-like individual, which suggests that the deletion affecting *SET* is very likely contributing to her phenotype. On the other hand, the association of likely pathogenic variants in NUPs is slowly emerging in individuals with NDDs (Fichtman et al., 2019). More specifically, haplo-insufficiency in NUP188 due to biallelic protein truncating variants has now been recognised to cause a new development disorder with a severe phenotype including microcephaly, trigonocephaly, congenital heart defects, congenital bilateral cataracts, microphthalmia and early death (Sandestig et al., 2019; C. A. Stevens & Krantz, 2018). Our RTT-like individual had both an N-terminal deletion of *NUP188* (from the large *de novo* deletion) and a heterozygous variant affecting a conserved residue on the other allele that had a low mean allele frequency in population databases. Our preliminary studies in fibroblasts from the RTT-like individual demonstrated reduced protein expression of NUP188 and enlarged nuclei (Theerthagiri et al., 2010). We hypothesise that the *NUP188* variant may have a lesser contribution to the individual's clinical features owing to her comparatively milder and distinct phenotype to that of other individuals with *NUP188* variants (Sandestig et al., 2019; Stevens & Krantz, 2018).

In addition to finding six RTT individuals with previously missed MECP2 variants, deep sequencing using the targeted sequencing panel of a classic RTT male enabled us to identify a mosaic variant in *MECP2* using DNA extracted from muscle (Chapter 6).

Variable mosaicism was observed in various tissues from the individual with the highest variant allele frequency in muscle and lowest in blood, the latter being the most common source of DNA tested in NGS studies. Thus, in this study we sequenced DNA from multiple tissues using high read-depth sequencing technologies with less stringent filtering methods in order to find the causative *MECP2* variants in this male with RTT features (Schonewolf-Greulich et al., 2019), further providing NGS data analysis strategies for the diagnosis of rare males with RTT/RTT-like features. In addition, these findings support the consideration of classic RTT as a clinical diagnosis when all the main criteria are met as specified in Neul's diagnostic criteria, in males as well as females. We suggest that low level mosaicism in *MECP2* should be tested in different tissues in classical RTT individuals before considering further genetic evaluations.

In summary, this project has identified pathogenic variations in four known RTT-related genes (*MECP2, SHANK3, SMC1A* and *EEF1A2*) and has further expanded the genotypic landscape of RTT to include six additional genes (*KIF1A*, *KAT6A*, *CHD8*, *SET*, *KCTD16,* and *NUP188*). We suggest that analysis of these genes should also be considered during NGS data curation from genetically undiagnosed RTT/RTT-like individuals. Moreover, this project has further emphasised the powerful role of the affordable singleton NGS in combination with streamlined advanced bioinformatics analysis to identify likely causative genes in *MECP2* negative RTT/ RTT-like individuals. Last but not the least, this study has also highlighted the utility of web-based freely accessible platforms, such as GeneMatcher, to enable researchers and clinicians with interest in similar genes to be connected with each other, with an ultimate goal of establishing the potential role of the novel genes with Mendelian phenotypes to further enhance our understanding of rare genetic disorders including RTT.

8.2 Future Directions

8.2.1 KIF1A

This study has demonstrated that individuals with *KIF1A* variants demonstrate a similar *in vitro* functional phenotype consisting of a significant reduction in intracellular trafficking of

KIF1A, measured by neurite tip accumulation assays and *in vitro* microtubule gliding assays. Although the structure and function of KIF1A has been well-established, currently there are no therapeutic agents for individuals with *KIF1A* variants. Moreover, to our knowledge, there is no active drug screening program to correct for impaired KIF1A function. To progress this research, we will aim to perform high throughput drug screening using a FDA-approved drug library to identify hit compounds for the improvement of impaired neurite based trafficking in a human cell model expressing wild type or variant KIF1A protein tagged with a fluorescent reporter. Following the identification of the top 10 compounds that increase the mobility of KIF1A, we will then determine whether these molecules functionally improve KIF1A in individual-specific neuronal cell lines with the individual variants p.(Asp248Glu) and p.(Pro305Leu) along with their gene-corrected isogenic controls that have already been established. We will evaluate a number of parameters including neuronal cell morphology (cell size, shape and neurite extension) using fixed cells stained with antibodies of interest and KIF1A specific cargo trafficking of fluorescently labelled BDNF, mitochondria and synaptotagmin, using already established methods (Meijering et al., 2004; Sbalzarini & Koumoutsakos, 2005; Shahen et al., 2018).

8.2.2 KAT6A

MeCP2 is a transcriptional regulator that associates with HDACs and silences genes by deacetylation of histone tails. Thus, in classic RTT individual reduced association of HDAC with chromatin due to MeCP2 deficiency would result in increased acetylation of the histone tail, resulting in transcriptional activation. However, KAT6A is a transcriptional activator that results in increased gene expression due to increased acetylation. Thus, defects in *KAT6A* would also result in aberrant gene expression, but further functional studies are needed to determine whether some variants cause their effect through a loss of function mechanism (resulting in reduced activation of key genes) or through a gain of function mechanism (resulting in over activation of genes). Our study has expanded the genetic landscape of RTT to include several genes including *KAT6A*, which encodes a HAT protein critical for transcriptional regulation through histone modifications. Previous studies have established the mechanistic link between such histone modifications and DNA methylation, which were then found to be associated with specific methylation signatures

on genomes of the affected individuals, further providing an alternative tool for diagnostic testing in individuals with NDDs (Aref-Eshghi et al., 2018). In future studies, we aim to perform genome-wide methylation analysis on the DNA samples from the seven RTT/RTT-like individuals in our cohort with pathogenic *KAT6A* variants. This may reveal key specific DNA methylation signatures that may aid in providing diagnosis for additional mutation-negative RTT/ RTT-like individuals.

8.2.3 *SET* and *NUP188*

During this project, we identified a RTT-like female with a large *de novo* deletion spanning 13 genes with two likely pathogenic genes *SET* and *NUP188* likely contributing towards her phenotype. Pathogenic *de novo* heterozygous missense and protein truncation variants have been previously reported in *SET* in individuals with an overlapping phenotype with our reported RTT individual, suggesting *SET* as the major contributing gene for her RTT phenotype. In order to validate these findings with functional evidence, we tested two different anti-SET antibodies; ab1183 (Abcam) and GTX 113834 (GeneTex). However, these antibodies were non-specific with multiple non-specific protein bands despite optimisation. Further plans involve looking for alternative anti-SET antibodies to evaluate SET protein expression levels in fibroblast from affected individual. With regard to the *NUP188* variant, our preliminary studies confirmed reduced protein expression levels and enlarged nuclei in the proband's fibroblasts, consistent with previous studies (Theerthagiri et al., 2010). Moreover, from our preliminary data, we have shown a subtle increase in the nuclear to cytoplasmic intensity ratio of polyA(+) RNA in RNA-FISH studies, suggesting total mRNA accumulation in the nucleus. Now that the experimental conditions have been optimised, we will be repeating these experiments with a full sample set. Furthermore, previous studies have shown an increased number of NPCs (a nuclear gateway made up of several NUP proteins) per nucleus (or per μm^2) in affected individual fibroblasts with loss of function *NUP188* variants (Theerthagiri et al., 2010). We are currently developing tools to measure the density and size of NPCs, and to detect any "central plug" phenotype depicting blocked NPCs due to trapped cargo, which may hinder transport of mRNA and proteins across the nuclear membrane.

8.2.4 KCTD16

We have uncovered a previously unrecognized link between KCTD16, critical element of GABAergic signalling pathway, and atypical RTT. Previous studies have demonstrated

that even a subtle change in GABA neurotransmission delay results in a severe clinical phenotype, often exhibited by individuals with NDDs, including RTT. Our preliminary functional studies of the homozygous *KCTD16* variant in *Xenopus* oocytes only showed a very subtle statistically insignificant change in electrical activity. Previous reports have used different read-outs based on the model system used (Schwenk et al., 2010), and so future research should use these approaches to further evaluate the pathogenicity of the *KCTD16* variant we have identified. Thus, we plan to re-evaluate the variant that we have identified using a more suitable mammalian CHO cell-based model for electrophysiological studies to measure the effect of the *KCTD16* variant on GABA-B receptor kinetics. CHO cells differ in their G-protein signalling pathway, and provide an alternate background for electrical studies without issues relating to the delayed solution exchange that are often observed in oocyte models.

8.2.5 Re-analysis of existing data and NGS of the remaining RTT/ RTT-like individuals

We will re-interrogate the NGS data to look for other potential variants that may have been missed during the first pass of data analysis, using less stringent variant filtering techniques and in the light of more recent novel gene discoveries. In addition, we will screen for copy number variants using the tool, Ximmer (Sadedin et al., 2018). Lastly, as 46 mutation-negative RTT/ RTT-like individuals in this cohort were not screened by NGS, we aim to undertake these investigations in the future.

In summary, this PhD project highlighted the powerful role of singleton NGS technologies in early diagnosis of mutation-negative RTT/ RTT-like individuals, ultimately putting an end to their often prolonged diagnostic odyssey, restoring reproductive confidence for families and enabling them to join a relevant supportive community, such as the Rett syndrome Association, KIF1A.org, and the KAT6A Foundation. In addition, this study has expanded the genotypic spectrum of RTT and phenotypic landscape of NDDs to include RTT/ RTT-like features through a combination of genomic sequencing, functional genomics and detailed phenotypic analysis of genetically undiagnosed individuals. Last but not the least, the findings from this study may be used to develop novel targeted therapies for at least some individuals with this devastating neurological disorder, for which there is currently no cure.

Bibliography

Bibliography

Abramoff, M. D., Magelhaes, P. J., & Ram, S. J. (2004). Image Processing with ImageJ. *Biophotonics International, 11*, 36-42.

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

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

Arafat, A., Jing, P., Ma, Y., Pu, M., Nan, G., Fang, H., . . . Fei, Y. (2017). Unexplained Early Infantile Epileptic Encephalopathy in Han Chinese Children: Next-Generation Sequencing and Phenotype Enriching. *Sci Rep, 7*, 46227. doi:10.1038/srep46227

Archer, H. L., Evans, J. C., Millar, D. S., Thompson, P. W., Kerr, A. M., Leonard, H., . . . Clarke, A. J. (2006). NTNG1 mutations are a rare cause of Rett syndrome. *Am J Med Genet A, 140*(7), 691-694. doi:10.1002/ajmg.a.31133

Aref-Eshghi, E., Rodenhiser, D. I., Schenkel, L. C., Lin, H., Skinner, C., Ainsworth, P., . . . Sadikovic, B. (2018). Genomic DNA Methylation Signatures Enable Concurrent Diagnosis and Clinical Genetic Variant Classification in Neurodevelopmental Syndromes. *Am J Hum Genet, 102*(1), 156-174. doi:10.1016/j.ajhg.2017.12.008

Ariani, F., Hayek, G., Rondinella, D., Artuso, R., Mencarelli, M. A., Spanhol-Rosseto, A., . . . Renieri, A. (2008). FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet, 83*(1), 89-93. doi:10.1016/j.ajhg.2008.05.015

Armstrong, A. H., Hangauer, J., Agazzi, H., Nunez, A., & Gieron-Korthals, M. (2010). Individuals with intellectual and developmental disabilities. In *Handbook of Pediatric Neuropsychology* (pp. 537-550). New York: Springer.

Armstrong, D. D. (2002). Neuropathology of Rett syndrome. *Ment Retard Dev Disabil Res Rev, 8*(2), 72-76. doi:10.1002/mrdd.10027

Armstrong, D. D., Dunn, J. K., Schultz, R. J., Herbert, D. A., Glaze, D. G., & Motil, K. J. (1999). Organ growth in Rett syndrome: a postmortem examination analysis. *Pediatr Neurol, 20*(2), 125-129. doi:10.1016/s0887-8994(98)00124-6

Ausio, J., Martinez de Paz, A., & Esteller, M. (2014). MeCP2: the long trip from a chromatin protein to neurological disorders. *Trends Mol Med, 20*(9), 487-498. doi:10.1016/j.molmed.2014.03.004

Ayhan, F., & Konopka, G. (2019). Regulatory genes and pathways disrupted in autism spectrum disorders. *Prog Neuropsychopharmacol Biol Psychiatry, 89*, 57-64. doi:10.1016/j.pnpbp.2018.08.017

Baasch, A. L., Huning, I., Gilissen, C., Klepper, J., Veltman, J. A., Gillessen-Kaesbach, G., . . . Lohmann, K. (2014). Exome sequencing identifies a de novo SCN2A mutation in a patient with intractable seizures, severe intellectual disability, optic atrophy, muscular hypotonia, and brain abnormalities. *Epilepsia, 55*(4), e25-29. doi:10.1111/epi.12554

Baltussen, L. L., Negraes, P. D., Silvestre, M., Claxton, S., Moeskops, M., Christodoulou, E., . . . Ultanir, S. K. (2018). Chemical genetic identification of CDKL5 substrates reveals its role in neuronal microtubule dynamics. *EMBO J, 37*(24). doi:10.15252/embj.201899763

Barbiero, I., De Rosa, R., & Kilstrup-Nielsen, C. (2019). Microtubules: A Key to Understand and Correct Neuronal Defects in CDKL5 Deficiency Disorder? *Int J Mol Sci, 20*(17), E4075. doi:10.3390/ijms20174075

Barbiero, I., Peroni, D., Siniscalchi, P., Rusconi, L., Tramarin, M., De Rosa, R., . . . Kilstrup-Nielsen, C. (2020). Pregnenolone and pregnenolone-methyl-ether rescue neuronal defects caused by dysfunctional CLIP170 in a neuronal model of CDKL5 Deficiency Disorder. *Neuropharmacology, 164*, 107897. doi:10.1016/j.neuropharm.2019.107897

Barbiero, I., Peroni, D., Tramarin, M., Chandola, C., Rusconi, L., Landsberger, N., & Kilstrup-Nielsen, C. (2017). The neurosteroid pregnenolone reverts microtubule derangement induced by the loss of a functional CDKL5-IQGAP1 complex. *Hum Mol Genet, 26*(18), 3520-3530. doi:10.1093/hmg/ddx237

Barnard, R. A., Pomaville, M. B., & O'Roak, B. J. (2015). Mutations and Modeling of the Chromatin Remodeler CHD8 Define an Emerging Autism Etiology. *Front Neurosci, 9*, 477. doi:10.3389/fnins.2015.00477

Bauman, M. L., Kemper, T. L., & Arin, D. M. (1995). Pervasive neuroanatomic abnormalities of the brain in three cases of Rett's syndrome. *Neurology, 45*(8), 1581-1586. doi:10.1212/wnl.45.8.1581

Belichenko, P. V., Oldfors, A., Hagberg, B., & Dahlstrom, A. (1994). Rett syndrome: 3-D confocal microscopy of cortical pyramidal dendrites and afferents. *Neuroreport, 5*(12), 1509-1513.

Ben Zeev, B., Bebbington, A., Ho, G., Leonard, H., de Klerk, N., Gak, E., . . . Christodoulou, J. (2009). The common BDNF polymorphism may be a modifier of disease severity in Rett syndrome. *Neurology, 72*(14), 1242-1247. doi:10.1212/01.wnl.0000345664.72220.6a

Bernier, R., Golzio, C., Xiong, B., Stessman, H. A., Coe, B. P., Penn, O., . . . Eichler, E. E. (2014). Disruptive CHD8 mutations define a subtype of autism early in development. *Cell, 158*(2), 263-276. doi:10.1016/j.cell.2014.06.017

Bienvenu, T., & Chelly, J. (2006). Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat Rev Genet, 7*(6), 415-426. doi:10.1038/nrg1878

Boyle, L., & Chung, W. (2017). *Expanding the natural history of KIF1A associated neurological disorders (KAND)*.

Boyle, L., & Chung, W. K. (2017, October 17-21). *Expanding the natural history of KIF1A associated neurological disorders (KAND).* Paper presented at the 67th Annual Meeting of The American Society of Human Genetics, Orlando, FL.

Bramswig, N. C., Ludecke, H. J., Alanay, Y., Albrecht, B., Barthelmie, A., Boduroglu, K., . . . Wieczorek, D. (2015). Exome sequencing unravels unexpected differential diagnoses in individuals with the tentative diagnosis of Coffin-Siris and Nicolaides-Baraitser syndromes. *Hum Genet, 134*(6), 553-568. doi:10.1007/s00439-015-1535-8

Buschdorf, J. P., & Stratling, W. H. (2004). A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. *J Mol Med (Berl), 82*(2), 135-143. doi:10.1007/s00109-003-0497-9

Byrd, A. K., & Raney, K. D. (2012). Superfamily 2 helicases. *Front Biosci (Landmark Ed), 17*, 2070-2088. doi:10.2741/4038

Campeau, P. M., Kim, J. C., Lu, J. T., Schwartzentruber, J. A., Abdul-Rahman, O. A., Schlaubitz, S., . . . Lee, B. H. (2012). Mutations in KAT6B, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet, 90*(2), 282-289. doi:10.1016/j.ajhg.2011.11.023

Cardoso, A. R., Lopes-Marques, M., Silva, R. M., Serrano, C., Amorim, A., Prata, M. J., & Azevedo, L. (2019). Essential genetic findings in neurodevelopmental disorders. *Hum Genomics, 13*(1), 31. doi:10.1186/s40246-019-0216-4

Carouge, D., Host, L., Aunis, D., Zwiller, J., & Anglard, P. (2010). CDKL5 is a brain MeCP2 target gene regulated by DNA methylation. *Neurobiol Dis, 38*(3), 414-424. doi:10.1016/j.nbd.2010.02.014

Cellot, G., & Cherubini, E. (2014). GABAergic signaling as therapeutic target for autism spectrum disorders. *Front Pediatr, 2*, 70. doi:10.3389/fped.2014.00070

Cervoni, N., Detich, N., Seo, S. B., Chakravarti, D., & Szyf, M. (2002). The oncoprotein Set/TAF-1beta, an inhibitor of histone acetyltransferase, inhibits active demethylation of DNA, integrating DNA methylation and transcriptional silencing. *J Biol Chem, 277*(28), 25026-25031. doi:10.1074/jbc.M202256200

Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T., Qin, J., & Zoghbi, H. Y. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science, 320*(5880), 1224-1229. doi:10.1126/science.1153252

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

Chao, H. T., Zoghbi, H. Y., & Rosenmund, C. (2007). MeCP2 controls excitatory synaptic strength by regulating glutamatergic synapse number. *Neuron, 56*(1), 58-65. doi:10.1016/j.neuron.2007.08.018

Chattopadhyaya, B., & Cristo, G. D. (2012). GABAergic circuit dysfunctions in neurodevelopmental disorders. *Front Psychiatry, 3*, 51. doi:10.3389/fpsyt.2012.00051

Chen, Q., Zhu, Y. C., Yu, J., Miao, S., Zheng, J., Xu, L., . . . Xiong, Z. Q. (2010). CDKL5, a protein associated with rett syndrome, regulates neuronal morphogenesis via Rac1 signaling. *J Neurosci, 30*(38), 12777-12786. doi:10.1523/JNEUROSCI.1102-10.2010

Cheng, J., Huang, M., Zhu, Y., Xin, Y. J., Zhao, Y. K., Huang, J., . . . Qiu, Z. (2014). SUMOylation of MeCP2 is essential for transcriptional repression and hippocampal synapse development. *J Neurochem, 128*(6), 798-806. doi:10.1111/jnc.12523

Cheon, C. K., Lim, S. H., Kim, Y. M., Kim, D., Lee, N. Y., Yoon, T. S., . . . Lee, J. R. (2017). Autosomal dominant transmission of complicated hereditary spastic paraplegia due to

a dominant negative mutation of KIF1A, SPG30 gene. *Sci Rep, 7*(1), 12527. doi:10.1038/s41598-017-12999-9

Christodoulou, J., Grimm, A., Maher, T., & Bennetts, B. (2003). RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat, 21*(5), 466-472.

Citterio, A., Arnoldi, A., Panzeri, E., Merlini, L., D'Angelo, M. G., Musumeci, O., . . . Bassi, M. T. (2015). Variants in KIF1A gene in dominant and sporadic forms of hereditary spastic paraparesis. *J Neurol, 262*(12), 2684-2690. doi:10.1007/s00415-015-7899-9

Clayton-Smith, J., O'Sullivan, J., Daly, S., Bhaskar, S., Day, R., Anderson, B., . . . Black, G. (2011). Whole-exome-sequencing identifies mutations in histone acetyltransferase gene KAT6B in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet, 89*(5), 675-681. doi:10.1016/j.ajhg.2011.10.008

Clayton-Smith, J., Watson, P., Ramsden, S., & Black, G. C. (2000). Somatic mutation in MECP2 as a non-fatal neurodevelopmental disorder in males. *Lancet, 356*(9232), 830-832. doi:10.1016/s0140-6736(00)02661-1

Cogliati, F., Giorgini, V., Masciadri, M., Bonati, M. T., Marchi, M., Cracco, I., . . . Russo, S. (2019). Pathogenic Variants in STXBP1 and in Genes for GABAa Receptor Subunities Cause Atypical Rett/Rett-like Phenotypes. *Int J Mol Sci, 20*(15), E3621. doi:10.3390/ijms20153621

Cohen, S., Gabel, H. W., Hemberg, M., Hutchinson, A. N., Sadacca, L. A., Ebert, D. H., . . . Greenberg, M. E. (2011). Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. *Neuron, 72*(1), 72-85. doi:10.1016/j.neuron.2011.08.022

Compagnone, N. A., Zhang, P., Vigne, J. L., & Mellon, S. H. (2000). Novel role for the nuclear phosphoprotein SET in transcriptional activation of P450c17 and initiation of neurosteroidogenesis. *Mol Endocrinol, 14*(6), 875-888. doi:10.1210/mend.14.6.0469

Cooper, G. (2000). Microtubules. In *The Cell: A Molecular Approach* (2nd ed.). Sunderland (MA): Sinauer Associates.

Cotney, J., Muhle, R. A., Sanders, S. J., Liu, L., Willsey, A. J., Niu, W., . . . Noonan, J. P. (2015). The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat Commun, 6*, 6404. doi:10.1038/ncomms7404

D'Angelo, M. A., Gomez-Cavazos, J. S., Mei, A., Lackner, D. H., & Hetzer, M. W. (2012). A change in nuclear pore complex composition regulates cell differentiation. *Dev Cell, 22*(2), 446-458. doi:10.1016/j.devcel.2011.11.021

Dani, V. S., Chang, Q., Maffei, A., Turrigiano, G. G., Jaenisch, R., & Nelson, S. B. (2005). Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. *Proc Natl Acad Sci U S A, 102*(35), 12560-12565. doi:10.1073/pnas.0506071102

Dani, V. S., & Nelson, S. B. (2009). Intact long-term potentiation but reduced connectivity between neocortical layer 5 pyramidal neurons in a mouse model of Rett syndrome. *J Neurosci, 29*(36), 11263-11270. doi:10.1523/JNEUROSCI.1019-09.2009

Dastidar, S. G., Narayanan, S., Stifani, S., & D'Mello, S. R. (2012). Transducin-like enhancer of Split-1 (TLE1) combines with Forkhead box protein G1 (FoxG1) to promote neuronal survival. *J Biol Chem, 287*(18), 14749-14759. doi:10.1074/jbc.M111.328336

De Bona, C., Zappella, M., Hayek, G., Meloni, I., Vitelli, F., Bruttini, M., . . . Renieri, A. (2000). Preserved speech variant is allelic of classic Rett syndrome. *Eur J Hum Genet, 8*(5), 325-330. doi:10.1038/sj.ejhg.5200473

De Felice, C., Maffei, S., Signorini, C., Leoncini, S., Lunghetti, S., Valacchi, G., . . . Hayek, J. (2012). Subclinical myocardial dysfunction in Rett syndrome. *Eur Heart J Cardiovasc Imaging, 13*(4), 339-345. doi:10.1093/ejechocard/jer256

De La Fuente, R., Baumann, C., & Viveiros, M. M. (2011). Role of ATRX in chromatin structure and function: implications for chromosome instability and human disease. *Reproduction, 142*(2), 221-234. doi:10.1530/REP-10-0380

de Ligt, J., Willemsen, M. H., van Bon, B. W., Kleefstra, T., Yntema, H. G., Kroes, T., . . . Vissers, L. E. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med, 367*(20), 1921-1929. doi:10.1056/NEJMoa1206524

De Rubeis, S., Siper, P. M., Durkin, A., Weissman, J., Muratet, F., Halpern, D., . . . Kolevzon, A. (2018). Delineation of the genetic and clinical spectrum of Phelan-McDermid syndrome caused by SHANK3 point mutations. *Mol Autism, 9*, 31. doi:10.1186/s13229-018-0205-9

Deciphering Developmental Disorders, S. (2017). Prevalence and architecture of de novo mutations in developmental disorders. *Nature, 542*(7642), 433-438. doi:10.1038/nature21062

Delepine, C., Meziane, H., Nectoux, J., Opitz, M., Smith, A. B., Ballatore, C., . . . Bienvenu, T. (2016). Altered microtubule dynamics and vesicular transport in mouse and human MeCP2-deficient astrocytes. *Hum Mol Genet, 25*(1), 146-157. doi:10.1093/hmg/ddv464

Delepine, C., Nectoux, J., Bahi-Buisson, N., Chelly, J., & Bienvenu, T. (2013). MeCP2 deficiency is associated with impaired microtubule stability. *FEBS Lett, 587*(2), 245-253. doi:10.1016/j.febslet.2012.11.033

Della Ragione, F., Filosa, S., Scalabri, F., & D'Esposito, M. (2012). MeCP2 as a genome-wide modulator: the renewal of an old story. *Front Genet, 3*, 181. doi:10.3389/fgene.2012.00181

Della Ragione, F., Vacca, M., Fioriniello, S., Pepe, G., & D'Esposito, M. (2016). MECP2, a multi-talented modulator of chromatin architecture. *Brief Funct Genomics, 15*(6), 420-431. doi:10.1093/bfgp/elw023

Di Cristo, G., Pizzorusso, T., Cancedda, L., & Sernagor, E. (2011). GABAergic circuit development and its implication for CNS disorders. *Neural Plast, 2011*, 623705. doi:10.1155/2011/623705

Du, Q., Luu, P. L., Stirzaker, C., & Clark, S. J. (2015). Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics, 7*(6), 1051-1073. doi:10.2217/epi.15.39

Ehrhart, F., Sangani, N. B., & Curfs, L. M. G. (2018). Current developments in the genetics of Rett and Rett-like syndrome. *Curr Opin Psychiatry, 31*(2), 103-108. doi:10.1097/YCO.0000000000000389

Epi K consortium, & Epilepsy Phenome/Genome Project. (2017). Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. *Lancet Neurol, 16*(2), 135-143. doi:10.1016/S1474-4422(16)30359-3

Erlich, Y., Edvardson, S., Hodges, E., Zenvirt, S., Thekkat, P., Shaag, A., . . . Elpeleg, O. (2011). Exome sequencing and disease-network analysis of a single family implicate a mutation in KIF1A in hereditary spastic paraparesis. *Genome Res, 21*(5), 658-664. doi:10.1101/gr.117143.110

Esmaeeli Nieh, S., Madou, M. R., Sirajuddin, M., Fregeau, B., McKnight, D., Lexa, K., . . . Sherr, E. H. (2015). De novo mutations in KIF1A cause progressive encephalopathy and brain atrophy. *Ann Clin Transl Neurol, 2*(6), 623-635. doi:10.1002/acn3.198

Evans, J. C., Archer, H. L., Colley, J. P., Ravn, K., Nielsen, J. B., Kerr, A., . . . Clarke, A. J. (2005). Early onset seizures and Rett-like features associated with mutations in CDKL5. *Eur J Hum Genet, 13*(10), 1113-1120. doi:10.1038/sj.ejhg.5201451

Fakhro, K. A., Choi, M., Ware, S. M., Belmont, J. W., Towbin, J. A., Lifton, R. P., . . . Brueckner, M. (2011). Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc Natl Acad Sci U S A, 108*(7), 2915-2920. doi:10.1073/pnas.1019645108

Fehr, S., Bebbington, A., Nassar, N., Downs, J., Ronen, G. M., N, D. E. K., & Leonard, H. (2011). Trends in the diagnosis of Rett syndrome in Australia. *Pediatr Res, 70*(3), 313-319. doi:10.1203/PDR.0b013e3182242461

Fehr, S., Wilson, M., Downs, J., Williams, S., Murgia, A., Sartori, S., . . . Christodoulou, J. (2013). The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *Eur J Hum Genet, 21*(3), 266-273. doi:10.1038/ejhg.2012.156

Fichtman, B., Harel, T., Biran, N., Zagairy, F., Applegate, C. D., Salzberg, Y., . . . Edvardson, S. (2019). Pathogenic Variants in NUP214 Cause "Plugged" Nuclear Pore Channels and Acute Febrile Encephalopathy. *Am J Hum Genet*. doi:10.1016/j.ajhg.2019.05.003

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., & den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat, 32*(5), 557-563. doi:10.1002/humu.21438

Fritzius, T., & Bettler, B. (2019). The organizing principle of GABAB receptor complexes: Physiological and pharmacological implications. *Basic Clin Pharmacol Toxicol*. doi:10.1111/bcpt.13241

Fritzius, T., Turecek, R., Seddik, R., Kobayashi, H., Tiao, J., Rem, P. D., . . . Bettler, B. (2017). KCTD Hetero-oligomers Confer Unique Kinetic Properties on Hippocampal GABAB Receptor-Induced K+ Currents. *J Neurosci, 37*(5), 1162-1175. doi:10.1523/JNEUROSCI.2181-16.2016

Fromer, M., Moran, J. L., Chambert, K., Banks, E., Bergen, S. E., Ruderfer, D. M., . . . Purcell, S. M. (2012). Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet, 91*(4), 597-607. doi:10.1016/j.ajhg.2012.08.005

Fujii-Nakata, T., Ishimi, Y., Okuda, A., & Kikuchi, A. (1992). Functional analysis of nucleosome assembly protein, NAP-1. The negatively charged COOH-terminal region is not necessary for the intrinsic assembly activity. *J Biol Chem, 267*(29), 20980-20986.

Gabrych, D. R., Lau, V. Z., Niwa, S., & Silverman, M. A. (2019). Going Too Far Is the Same as Falling Short(dagger): Kinesin-3 Family Members in Hereditary Spastic Paraplegia. *Front Cell Neurosci, 13*, 419. doi:10.3389/fncel.2019.00419

Gamper, N., Stockand, J. D., & Shapiro, M. S. (2005). The use of Chinese hamster ovary (CHO) cells in the study of ion channels. *J Pharmacol Toxicol Methods, 51*(3), 177-185. doi:10.1016/j.vascn.2004.08.008

Garrett, R., & Grisham, C. (2005). Molecular motors. In *Biochemistry*. London: Thomson Brooks/Cole.

Gervais, L., van den Beek, M., Josserand, M., Salle, J., Stefanutti, M., Perdigoto, C. N., . . . Bardin, A. J. (2019). Stem Cell Proliferation Is Kept in Check by the Chromatin Regulators Kismet/CHD7/CHD8 and Trr/MLL3/4. *Dev Cell, 49*(4), 556-573 e556. doi:10.1016/j.devcel.2019.04.033

Gibson, J. H., Slobedman, B., K, N. H., Williamson, S. L., Minchenko, D., El-Osta, A., . . . Christodoulou, J. (2010). Downstream targets of methyl CpG binding protein 2 and their abnormal expression in the frontal cortex of the human Rett syndrome brain. *BMC Neurosci, 11*, 53. doi:10.1186/1471-2202-11-53

Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W., Willemsen, M. H., . . . Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature, 511*(7509), 344-347. doi:10.1038/nature13394

Glotzer, M. (2009). The 3Ms of central spindle assembly: microtubules, motors and MAPs. *Nat Rev Mol Cell Biol, 10*(1), 9-20. doi:10.1038/nrm2609

Godler, D. E., & Amor, D. J. (2019). DNA methylation analysis for screening and diagnostic testing in neurodevelopmental disorders. *Essays Biochem, 63*(6), 785-795. doi:10.1042/EBC20190056

Gold, W. A., Lacina, T. A., Cantrill, L. C., & Christodoulou, J. (2015). MeCP2 deficiency is associated with reduced levels of tubulin acetylation and can be restored using HDAC6 inhibitors. *J Mol Med (Berl), 93*(1), 63-72. doi:10.1007/s00109-014-1202-x

Gompers, A. L., Su-Feher, L., Ellegood, J., Copping, N. A., Riyadh, M. A., Stradleigh, T. W., . . . Nord, A. S. (2017). Germline Chd8 haploinsufficiency alters brain development in mouse. *Nat Neurosci, 20*(8), 1062-1073. doi:10.1038/nn.4592

Gonzales, M. L., Adams, S., Dunaway, K. W., & LaSalle, J. M. (2012). Phosphorylation of distinct sites in MeCP2 modifies cofactor associations and the dynamics of transcriptional regulation. *Mol Cell Biol, 32*(14), 2894-2903. doi:10.1128/MCB.06728-11

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. *Science, 185*(4154), 862-864. doi:10.1126/science.185.4154.862

Guang, S., Pang, N., Deng, X., Yang, L., He, F., Wu, L., . . . Peng, J. (2018). Synaptopathology Involved in Autism Spectrum Disorder. *Front Cell Neurosci, 12*, 470. doi:10.3389/fncel.2018.00470

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

Hagberg, B., & Rasmussen, P. (1986). "Forme fruste" of Rett syndrome--a case report. *Am J Med Genet Suppl, 1*, 175-181.

Hamdan, F. F., Gauthier, J., Araki, Y., Lin, D. T., Yoshizawa, Y., Higashi, K., . . . Michaud, J. L. (2011). Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. *Am J Hum Genet, 88*(3), 306-316. doi:10.1016/j.ajhg.2011.02.001

Hamdan, F. F., Srour, M., Capo-Chichi, J. M., Daoud, H., Nassif, C., Patry, L., . . . Michaud, J. L. (2014). De novo mutations in moderate or severe intellectual disability. *PLoS Genet, 10*(10), e1004772. doi:10.1371/journal.pgen.1004772

Hammond, J. W., Cai, D., Blasius, T. L., Li, Z., Jiang, Y., Jih, G. T., . . . Verhey, K. J. (2009). Mammalian Kinesin-3 motors are dimeric in vivo and move by processive motility upon release of autoinhibition. *PLoS Biol, 7*(3), e72. doi:10.1371/journal.pbio.1000072

Hanahan, D., Jessee, J., & Bloom, F. R. (1991). Plasmid transformation of Escherichia coli and other bacteria. *Methods Enzymol, 204*, 63-113. doi:10.1016/0076-6879(91)04006-a

Hanefeld, F. (1985). The clinical pattern of the Rett syndrome. *Brain Dev, 7*(3), 320-325.

Hara, M., Ohba, C., Yamashita, Y., Saitsu, H., Matsumoto, N., & Matsuishi, T. (2015). De novo SHANK3 mutation causes Rett syndrome-like phenotype in a female patient. *Am J Med Genet A, 167*(7), 1593-1596. doi:10.1002/ajmg.a.36775

Haskell, G. T., Jensen, B. C., Samsa, L. A., Marchuk, D., Huang, W., Skrzynia, C., . . . Berg, J. S. (2017). Whole Exome Sequencing Identifies Truncating Variants in Nuclear Envelope Genes in Patients With Cardiovascular Disease. *Circ Cardiovasc Genet, 10*(3). doi:10.1161/CIRCGENETICS.116.001443

Hector, R. D., Kalscheuer, V. M., Hennig, F., Leonard, H., Downs, J., Clarke, A., . . . Cobb, S. R. (2017). CDKL5 variants: Improving our understanding of a rare neurologic disorder. *Neurol Genet, 3*(6), e200. doi:10.1212/NXG.0000000000000200

Helander, A., Stodberg, T., Jaeken, J., Matthijs, G., Eriksson, M., & Eggertsen, G. (2013). Dolichol kinase deficiency (DOLK-CDG) with a purely neurological presentation caused by a novel mutation. *Mol Genet Metab, 110*(3), 342-344. doi:10.1016/j.ymgme.2013.07.002

Henriksen, M. W., Ravn, K., Paus, B., von Tetzchner, S., & Skjeldal, O. H. (2018). De novo mutations in SCN1A are associated with classic Rett syndrome: a case report. *BMC Med Genet, 19*(1), 184. doi:10.1186/s12881-018-0700-z

Hill, V. K., Kim, J. S., & Waldman, T. (2016). Cohesin mutations in human cancer. *Biochim Biophys Acta, 1866*(1), 1-11. doi:10.1016/j.bbcan.2016.05.002

Hirokawa, N., Noda, Y., Tanaka, Y., & Niwa, S. (2009). Kinesin superfamily motor proteins and intracellular transport. *Nat Rev Mol Cell Biol, 10*(10), 682-696. doi:10.1038/nrm2774

Holder, J. L., Jr., & Quach, M. M. (2016). The spectrum of epilepsy and electroencephalographic abnormalities due to SHANK3 loss-of-function mutations. *Epilepsia, 57*(10), 1651-1659. doi:10.1111/epi.13506

Hota, S. K., & Bruneau, B. G. (2016). ATP-dependent chromatin remodeling during mammalian development. *Development, 143*(16), 2882-2897. doi:10.1242/dev.128892

Hotchkiss, L., Donkervoort, S., Leach, M. E., Mohassel, P., Bharucha-Goebel, D. X., Bradley, N., . . . Bonnemann, C. G. (2016). Novel De Novo Mutations in KIF1A as a Cause of Hereditary Spastic Paraplegia With Progressive Central Nervous System Involvement. *J Child Neurol, 31*(9), 1114-1119. doi:10.1177/0883073816639718

Hsu, P., Ma, A., Wilson, M., Williams, G., Curotta, J., Munns, C. F., & Mehr, S. (2014). CHARGE syndrome: a review. *J Paediatr Child Health, 50*(7), 504-511. doi:10.1111/jpc.12497

Huang, C. F., & Banker, G. (2012). The translocation selectivity of the kinesins that mediate neuronal organelle transport. *Traffic, 13*(4), 549-564. doi:10.1111/j.1600-0854.2011.01325.x

Hughes, V. (2012). SHANK3 mice comparisons reveal array of differences. *Spectrum*. Retrieved from https://www.spectrumnews.org/news/shank3-mice-comparisons-reveal-array-of-differences/

Huisman, S., Mulder, P. A., Redeker, E., Bader, I., Bisgaard, A. M., Brooks, A., . . . Hennekam, R. C. (2017). Phenotypes and genotypes in individuals with SMC1A variants. *Am J Med Genet A, 173*(8), 2108-2125. doi:10.1002/ajmg.a.38279

Inui, T., Kobayashi, S., Ashikari, Y., Sato, R., Endo, W., Uematsu, M., . . . Haginoya, K. (2016). Two cases of early-onset myoclonic seizures with continuous parietal delta activity caused by EEF1A2 mutations. *Brain Dev, 38*(5), 520-524. doi:10.1016/j.braindev.2015.11.003

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., . . . Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature, 515*(7526), 216-221. doi:10.1038/nature13908

Ishiguro, K. I. (2019). The cohesin complex in mammalian meiosis. *Genes Cells, 24*(1), 6-30. doi:10.1111/gtc.12652

Iwama, K., Mizuguchi, T., Takeshita, E., Nakagawa, E., Okazaki, T., Nomura, Y., . . . Matsumoto, N. (2019). Genetic landscape of Rett syndrome-like phenotypes revealed by whole exome sequencing. *J Med Genet, 56*(6), 396-407. doi:10.1136/jmedgenet-2018-105775

Iwase, S., & Martin, D. M. (2018). Chromatin in nervous system development and disease. *Mol Cell Neurosci, 87*, 1-3. doi:10.1016/j.mcn.2017.12.006

Jang, D. H., Chae, H., & Kim, M. (2015). Autistic and Rett-like features associated with 2q33.3-q34 interstitial deletion. *Am J Med Genet A, 167A*(9), 2213-2218. doi:10.1002/ajmg.a.37119

Jessberger, R., Podust, V., Hubscher, U., & Berg, P. (1993). A mammalian protein complex that repairs double-strand breaks and deletions by recombination. *J Biol Chem, 268*(20), 15070-15079.

Jessberger, R., Riwar, B., Baechtold, H., & Akhmedov, A. T. (1996). SMC proteins constitute two subunits of the mammalian recombination complex RC-1. *EMBO J, 15*(15), 4061-4068.

Jian, L., Nagarajan, L., de Klerk, N., Ravine, D., Bower, C., Anderson, A., . . . Leonard, H. (2006). Predictors of seizure onset in Rett syndrome. *J Pediatr, 149*(4), 542-547. doi:10.1016/j.jpeds.2006.06.015

Jian, L., Nagarajan, L., de Klerk, N., Ravine, D., Christodoulou, J., & Leonard, H. (2007). Seizures in Rett syndrome: an overview from a one-year calendar study. *Eur J Paediatr Neurol, 11*(5), 310-317. doi:10.1016/j.ejpn.2007.02.008

Jiang, Y., Oldridge, D. A., Diskin, S. J., & Zhang, N. R. (2015). CODEX: a normalization and copy number variation detection method for whole exome sequencing. *Nucleic Acids Res, 43*(6), e39. doi:10.1093/nar/gku1363

Jordan, C., Li, H. H., Kwan, H. C., & Francke, U. (2007). Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. *BMC Med Genet, 8*, 36. doi:10.1186/1471-2350-8-36

Joseph, N., Hutterer, A., Poser, I., & Mishima, M. (2012). ARF6 GTPase protects the post-mitotic midbody from 14-3-3-mediated disintegration. *EMBO J, 31*(11), 2604-2614. doi:10.1038/emboj.2012.139

Kalscheuer, V. M., Tao, J., Donnelly, A., Hollway, G., Schwinger, E., Kubart, S., . . . Gecz, J. (2003). Disruption of the serine/threonine kinase 9 gene causes severe X-linked infantile spasms and mental retardation. *Am J Hum Genet, 72*(6), 1401-1411. doi:10.1086/375538

Karaca, E., Harel, T., Pehlivan, D., Jhangiani, S. N., Gambin, T., Coban Akdemir, Z., . . . Lupski, J. R. (2015). Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. *Neuron, 88*(3), 499-513. doi:10.1016/j.neuron.2015.09.048

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., . . . MacArthur, D. G. (2019). Variation across 141,456 human exomes and genomes

reveals the spectrum of loss-of-function intolerance across human protein-coding genes. 531210. doi:10.1101/531210 %J bioRxiv

Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., . . . Nakayama, K. I. (2016). CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature, 537*(7622), 675-679. doi:10.1038/nature19357

Kaufmann, W. E., Naidu, S., & Budden, S. (1995). Abnormal expression of microtubule-associated protein 2 (MAP-2) in neocortex in Rett syndrome. *Neuropediatrics, 26*(2), 109-113. doi:10.1055/s-2007-979738

Kaur, S., Van Bergen, N. J., Gold, W. A., Eggers, S., Lunke, S., White, S. M., . . . Christodoulou, J. (2019). Whole exome sequencing reveals a de novo missense variant in EEF1A2 in a Rett syndrome-like patient. *Clin Case Rep, 7*(12), 2476-2482. doi:10.1002/ccr3.2511

Kennedy, J., Goudie, D., Blair, E., Chandler, K., Joss, S., McKay, V., . . . Newbury-Ecob, R. (2019). KAT6A Syndrome: genotype-phenotype correlation in 76 patients with pathogenic KAT6A variants. *Genet Med, 21*(4), 850-860. doi:10.1038/s41436-018-0259-2

Kikkawa, M., Sablin, E. P., Okada, Y., Yajima, H., Fletterick, R. J., & Hirokawa, N. (2001). Switch-based mechanism of kinesin motors. *Nature, 411*(6836), 439-445. doi:10.1038/35078000

Kim, Y. R., Park, J. B., Lee, Y. J., Hong, M. J., Kim, H. T., & Kim, H. J. (2017). Identifying the KAT6B Mutation via Diagnostic Exome Sequencing to Diagnose Say-Barber-Biesecker-Young-Simpson Syndrome in Three Generations of a Family. *Ann Rehabil Med, 41*(3), 505-510. doi:10.5535/arm.2017.41.3.505

Kishi, N., & Macklis, J. D. (2004). MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. *Mol Cell Neurosci, 27*(3), 306-321. doi:10.1016/j.mcn.2004.07.006

Klebe, S., Lossos, A., Azzedine, H., Mundwiller, E., Sheffer, R., Gaussen, M., . . . Stevanin, G. (2012). KIF1A missense mutations in SPG30, an autosomal recessive spastic paraplegia: distinct phenotypes according to the nature of the mutations. *Eur J Hum Genet, 20*(6), 645-649. doi:10.1038/ejhg.2011.261

Kline, A. D., Moss, J. F., Selicorni, A., Bisgaard, A. M., Deardorff, M. A., Gillett, P. M., . . . Hennekam, R. C. (2018). Diagnosis and management of Cornelia de Lange syndrome:

first international consensus statement. *Nat Rev Genet, 19*(10), 649-666. doi:10.1038/s41576-018-0031-0

Kortum, F., Das, S., Flindt, M., Morris-Rosendahl, D. J., Stefanova, I., Goldstein, A., . . . Dobyns, W. B. (2011). The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *J Med Genet, 48*(6), 396-406. doi:10.1136/jmg.2010.087528

Kranz, C., Jungeblut, C., Denecke, J., Erlekotte, A., Sohlbach, C., Debus, V., . . . Marquardt, T. (2007). A defect in dolichol phosphate biosynthesis causes a new inherited disorder with death in early infancy. *Am J Hum Genet, 80*(3), 433-440. doi:10.1086/512130

Krenn, M., Zulehner, G., Hotzy, C., Rath, J., Stogmann, E., Wagner, M., . . . Zimprich, F. (2017). Hereditary spastic paraplegia caused by compound heterozygous mutations outside the motor domain of the KIF1A gene. *Eur J Neurol, 24*(5), 741-747. doi:10.1111/ene.13279

Kriaucionis, S., & Bird, A. (2004). The major form of MeCP2 has a novel N-terminus generated by alternative splicing. *Nucleic Acids Res, 32*(5), 1818-1823. doi:10.1093/nar/gkh349

Krishnaraj, R., Ho, G., & Christodoulou, J. (2017). RettBASE: Rett syndrome database update. *Hum Mutat, 38*(8), 922-931. doi:10.1002/humu.23263

Kull, F. J., & Endow, S. A. (2013). Force generation by kinesin and myosin cytoskeletal motor proteins. *J Cell Sci, 126*(Pt 1), 9-19. doi:10.1242/jcs.103911

Kumar, S., Reynolds, K., Ji, Y., Gu, R., Rai, S., & Zhou, C. J. (2019). Impaired neurodevelopmental pathways in autism spectrum disorder: a review of signaling mechanisms and crosstalk. *J Neurodev Disord, 11*(1), 10. doi:10.1186/s11689-019-9268-y

Lam, B. D., Anthony, E. C., & Hordijk, P. L. (2012). Analysis of nucleo-cytoplasmic shuttling of the proto-oncogene SET/I2PP2A. *Cytometry A, 81*(1), 81-89. doi:10.1002/cyto.a.21153

Lam, B. D., Anthony, E. C., & Hordijk, P. L. (2013). Cytoplasmic targeting of the proto-oncogene SET promotes cell spreading and migration. *FEBS Lett, 587*(2), 111-119. doi:10.1016/j.febslet.2012.11.013

Lam, W., Millichap, J., Soares, D., Chin, R., McLellan, A., FitzPatrick, D., . . . Abbott, C. (2016). Novel de novo EEF1A2 missense mutations causing epilepsy and intellectual disability. *Mol Genet Genomic Med, 4*(4), 465–474.

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., . . . Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res, 46*(D1), D1062-D1067. doi:10.1093/nar/gkx1153

Langlois, S., Tarailo-Graovac, M., Sayson, B., Drogemoller, B., Swenerton, A., Ross, C. J., . . . van Karnebeek, C. D. (2016). De novo dominant variants affecting the motor domain of KIF1A are a cause of PEHO syndrome. *Eur J Hum Genet, 24*(6), 949-953. doi:10.1038/ejhg.2015.217

Lasser, M., Tiber, J., & Lowery, L. A. (2018). The Role of the Microtubule Cytoskeleton in Neurodevelopmental Disorders. *Front Cell Neurosci, 12*, 165. doi:10.3389/fncel.2018.00165

Latcheva, N. K., Delaney, T. L., Viveiros, J. M., Smith, R. A., Bernard, K. M., Harsin, B., . . . Liebl, F. L. W. (2019). The CHD Protein, Kismet, is Important for the Recycling of Synaptic Vesicles during Endocytosis. *Sci Rep, 9*(1), 19368. doi:10.1038/s41598-019-55900-6

Lee, J. R., Srour, M., Kim, D., Hamdan, F. F., Lim, S. H., Brunel-Guitton, C., . . . Michaud, J. L. (2015). De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. *Hum Mutat, 36*(1), 69-78. doi:10.1002/humu.22709

Lee, J. S., Yoo, Y., Lim, B. C., Kim, K. J., Song, J., Choi, M., & Chae, J. H. (2016). GM3 synthase deficiency due to ST3GAL5 variants in two Korean female siblings: Masquerading as Rett syndrome-like phenotype. *Am J Med Genet A, 170*(8), 2200-2205. doi:10.1002/ajmg.a.37773

Lefeber, D. J., de Brouwer, A. P., Morava, E., Riemersma, M., Schuurs-Hoeijmakers, J. H., Absmanner, B., . . . Wevers, R. A. (2011). Autosomal recessive dilated cardiomyopathy due to DOLK mutations results from abnormal dystroglycan O-mannosylation. *PLoS Genet, 7*(12), e1002427. doi:10.1371/journal.pgen.1002427

Lelieveld, S. H., Reijnders, M. R., Pfundt, R., Yntema, H. G., Kamsteeg, E. J., de Vries, P., . . . Gilissen, C. (2016). Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci, 19*(9), 1194-1196. doi:10.1038/nn.4352

Leonard, H., Colvin, L., Christodoulou, J., Schiavello, T., Williamson, S., Davis, M., . . . Yamashita, Y. (2003). Patients with the R133C mutation: is their phenotype different from patients with Rett syndrome with other mutations? *J Med Genet, 40*(5), e52. doi:10.1136/jmg.40.5.e52

Li, M., Milligan, C. J., Wang, H., Walker, A., Churilov, L., Lawrence, A. J., . . . Petrou, S. (2017). KCTD12 modulation of GABA(B) receptor function. *Pharmacol Res Perspect, 5*(4), e00319. doi:10.1002/prp2.319

Li, W., Calfa, G., Larimore, J., & Pozzo-Miller, L. (2012). Activity-dependent BDNF release and TRPC signaling is impaired in hippocampal neurons of Mecp2 mutant mice. *Proc Natl Acad Sci U S A, 109*(42), 17087-17092. doi:10.1073/pnas.1205271109

Lieu, M. T., Ng, B. G., Rush, J. S., Wood, T., Basehore, M. J., Hegde, M., . . . Wang, R. Y. (2013). Severe, fatal multisystem manifestations in a patient with dolichol kinase-congenital disorder of glycosylation. *Mol Genet Metab, 110*(4), 484-489. doi:10.1016/j.ymgme.2013.09.016

Liman, E. R., Tytgat, J., & Hess, P. (1992). Subunit stoichiometry of a mammalian K+ channel determined by construction of multimeric cDNAs. *Neuron, 9*(5), 861-871. doi:10.1016/0896-6273(92)90239-a

Lindeboom, R. G., Supek, F., & Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat Genet, 48*(10), 1112-1118. doi:10.1038/ng.3664

Liu, G. H., Guan, T., Datta, K., Coppinger, J., Yates, J., 3rd, & Gerace, L. (2009). Regulation of myoblast differentiation by the nuclear envelope protein NET39. *Mol Cell Biol, 29*(21), 5800-5812. doi:10.1128/MCB.00684-09

Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., . . . Ren, J. (2015). IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics, 31*(20), 3359-3361. doi:10.1093/bioinformatics/btv362

Liu, Z., Xiang, Y., & Sun, G. (2013). The KCTD family of proteins: structure, function, disease relevance. *Cell Biosci, 3*(1), 45. doi:10.1186/2045-3701-3-45

Lopes, F., Barbosa, M., Ameur, A., Soares, G., de Sa, J., Dias, A. I., . . . Maciel, P. (2016). Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet, 53*(3), 190-199. doi:10.1136/jmedgenet-2015-103568

Lundsgaard, M., Le, V. Q., Ernst, A., Laugaard-Jacobsen, H. C., Rasmussen, K., Pedersen, I. S., & Petersen, M. B. (2017). De novo KAT6B Mutation Identified with Whole-Exome

Sequencing in a Girl with Say-Barber/Biesecker/Young-Simpson Syndrome. *Mol Syndromol, 8*(1), 24-29. doi:10.1159/000452258

Maezawa, I., Swanberg, S., Harvey, D., LaSalle, J. M., & Jin, L. W. (2009). Rett syndrome astrocytes are abnormal and spread MeCP2 deficiency through gap junctions. *J Neurosci, 29*(16), 5051-5061. doi:10.1523/JNEUROSCI.0324-09.2009

Manning, B. J., & Yusufzai, T. (2017). The ATP-dependent chromatin remodeling enzymes CHD6, CHD7, and CHD8 exhibit distinct nucleosome binding and remodeling activities. *J Biol Chem, 292*(28), 11927-11936. doi:10.1074/jbc.M117.779470

Mannini, L., Liu, J., Krantz, I. D., & Musio, A. (2010). Spectrum and consequences of SMC1A mutations: the unexpected involvement of a core component of cohesin in human disease. *Hum Mutat, 31*(1), 5-10. doi:10.1002/humu.21129

Manuel, M. N., Martynoga, B., Molinek, M. D., Quinn, J. C., Kroemmer, C., Mason, J. O., & Price, D. J. (2011). The transcription factor Foxg1 regulates telencephalic progenitor proliferation cell autonomously, in part by controlling Pax6 expression levels. *Neural Dev, 6*, 9. doi:10.1186/1749-8104-6-9

Mari, F., Azimonti, S., Bertani, I., Bolognese, F., Colombo, E., Caselli, R., . . . Landsberger, N. (2005). CDKL5 belongs to the same molecular pathway of MeCP2 and it is responsible for the early-onset seizure variant of Rett syndrome. *Hum Mol Genet, 14*(14), 1935-1946. doi:10.1093/hmg/ddi198

Marschik, P. B., Einspieler, C., Oberle, A., Laccone, F., & Prechtl, H. F. (2009). Case report: Retracing atypical development: a preserved speech variant of Rett syndrome. *J Autism Dev Disord, 39*(6), 958-961. doi:10.1007/s10803-009-0703-x

Meijering, E., Jacob, M., Sarria, J. C., Steiner, P., Hirling, H., & Unser, M. (2004). Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry A, 58*(2), 167-176. doi:10.1002/cyto.a.20022

Miao, M., Ryan, K. J., & Wente, S. R. (2006). The integral membrane protein Pom34p functionally links nucleoporin subcomplexes. *Genetics, 172*(3), 1441-1457. doi:10.1534/genetics.105.052068

Miller, R. W., & Rubinstein, J. H. (1995). Tumors in Rubinstein-Taybi syndrome. *Am J Med Genet, 56*(1), 112-115. doi:10.1002/ajmg.1320560125

Mirabella, A. C., Foster, B. M., & Bartke, T. (2016). Chromatin deregulation in disease. *Chromosoma, 125*(1), 75-93. doi:10.1007/s00412-015-0530-0

Mitter, D., Pringsheim, M., Kaulisch, M., Plumacher, K. S., Schroder, S., Warthemann, R., . . . Brockmann, K. (2018). FOXG1 syndrome: genotype-phenotype association in 83 patients with FOXG1 variants. *Genet Med, 20*(1), 98-108. doi:10.1038/gim.2017.75

Miyoshi, G., & Fishell, G. (2012). Dynamic FoxG1 expression coordinates the integration of multipolar pyramidal neuron precursors into the cortical plate. *Neuron, 74*(6), 1045-1058. doi:10.1016/j.neuron.2012.04.025

Mullin, A. P., Gokhale, A., Moreno-De-Luca, A., Sanyal, S., Waddington, J. L., & Faundez, V. (2013). Neurodevelopmental disorders: mechanisms and boundary definitions from genomes, interactomes and proteomes. *Transl Psychiatry, 3*, e329. doi:10.1038/tp.2013.108

Nagarajan, R. P., Hogart, A. R., Gwye, Y., Martin, M. R., & LaSalle, J. M. (2006). Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. *Epigenetics, 1*(4), e1-11.

Nakajima, J., Okamoto, N., Tohyama, J., Kato, M., Arai, H., Funahashi, O., . . . Miyake, N. (2015). De novo EEF1A2 mutations in patients with characteristic facial features, intellectual disability, autistic behaviors and epilepsy. *Clin Genet, 87*(4), 356-361. doi:10.1111/cge.12394

Nawaz, M. S., Giarda, E., Bedogni, F., La Montanara, P., Ricciardi, S., Ciceri, D., . . . Kilstrup-Nielsen, C. (2016). CDKL5 and Shootin1 Interact and Concur in Regulating Neuronal Polarization. *PLoS One, 11*(2), e0148634. doi:10.1371/journal.pone.0148634

Nectoux, J., Florian, C., Delepine, C., Bahi-Buisson, N., Khelfaoui, M., Reibel, S., . . . Bienvenu, T. (2012). Altered microtubule dynamics in Mecp2-deficient astrocytes. *J Neurosci Res, 90*(5), 990-998. doi:10.1002/jnr.23001

Neul, J. L. (2012). The relationship of Rett syndrome and MECP2 disorders to autism. *Dialogues Clin Neurosci, 14*(3), 253-262.

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

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

Nikolic, M., Dudek, H., Kwon, Y. T., Ramos, Y. F., & Tsai, L. H. (1996). The cdk5/p35 kinase is essential for neurite outgrowth during neuronal differentiation. *Genes Dev, 10*(7), 816-825. doi:10.1101/gad.10.7.816

Nishiyama, M., Oshikawa, K., Tsukada, Y., Nakagawa, T., Iemura, S., Natsume, T., . . . Nakayama, K. I. (2009). CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nat Cell Biol, 11*(2), 172-182. doi:10.1038/ncb1831

Nitarska, J., Smith, J. G., Sherlock, W. T., Hillege, M. M., Nott, A., Barshop, W. D., . . . Riccio, A. (2016). A Functional Switch of NuRD Chromatin Remodeling Complex Subunits Regulates Mouse Cortical Development. *Cell Rep, 17*(6), 1683-1698. doi:10.1016/j.celrep.2016.10.022

Nitta, R., Kikkawa, M., Okada, Y., & Hirokawa, N. (2004). KIF1A alternately uses two loops to bind microtubules. *Science, 305*(5684), 678-683. doi:10.1126/science.1096621

Nomura, Y. (2005). Early behavior characteristics and sleep disturbance in Rett syndrome. *Brain Dev, 27 Suppl 1*, S35-S42. doi:10.1016/j.braindev.2005.03.017

Oexle, K., Thamm-Mucke, B., Mayer, T., & Tinschert, S. (2005). Macrocephalic mental retardation associated with a novel C-terminal MECP2 frameshift deletion. *Eur J Pediatr, 164*(3), 154-157. doi:10.1007/s00431-004-1583-x

Ohba, C., Haginoya, K., Osaka, H., Kubota, K., Ishiyama, A., Hiraide, T., . . . Matsumoto, N. (2015). De novo KIF1A mutations cause intellectual deficit, cerebellar atrophy, lower limb spasticity and visual disturbance. *J Hum Genet, 60*(12), 739-742. doi:10.1038/jhg.2015.108

Ohshima, T., Ward, J. M., Huh, C. G., Longenecker, G., Veeranna, Pant, H. C., . . . Kulkarni, A. B. (1996). Targeted disruption of the cyclin-dependent kinase 5 gene results in abnormal corticogenesis, neuronal pathology and perinatal death. *Proc Natl Acad Sci U S A, 93*(20), 11173-11178. doi:10.1073/pnas.93.20.11173

Okada, Y., Yamazaki, H., Sekine-Aizawa, Y., & Hirokawa, N. (1995). The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell, 81*(5), 769-780.

Okamoto, N., Miya, F., Tsunoda, T., Yanagihara, K., Kato, M., Saitoh, S., . . . Kosaki, K. (2014). KIF1A mutation in a patient with progressive neurodegeneration. *J Hum Genet, 59*(11), 639-641. doi:10.1038/jhg.2014.80

Olson, H. E., Demarest, S. T., Pestana-Knight, E. M., Swanson, L. C., Iqbal, S., Lal, D., . . . Benke, T. A. (2019). Cyclin-Dependent Kinase-Like 5 Deficiency Disorder: Clinical Review. *Pediatr Neurol, 97*, 18-25. doi:10.1016/j.pediatrneurol.2019.02.015

Padgett, C. L., & Slesinger, P. A. (2010). GABAB receptor coupling to G-proteins and ion channels. *Adv Pharmacol, 58*, 123-147. doi:10.1016/S1054-3589(10)58006-2

Park, Y. J., & Luger, K. (2006). Structure and function of nucleosome assembly proteins. *Biochem Cell Biol, 84*(4), 549-558. doi:10.1139/o06-088

Pelka, G. J., Watson, C. M., Radziewic, T., Hayward, M., Lahooti, H., Christodoulou, J., & Tam, P. P. (2006). Mecp2 deficiency is associated with learning and cognitive deficits and altered gene activity in the hippocampal region of mice. *Brain, 129*(Pt 4), 887-898. doi:10.1093/brain/awl022

Pennings, M., Schouten, M. I., van Gaalen, J., Meijer, R. P. P., de Bot, S. T., Kriek, M., . . . Kamsteeg, E. J. (2020). KIF1A variants are a frequent cause of autosomal dominant hereditary spastic paraplegia. *Eur J Hum Genet, 28*(1), 40-49. doi:10.1038/s41431-019-0497-z

Petrij, F., Giles, R. H., Dauwerse, H. G., Saris, J. J., Hennekam, R. C., Masuno, M., . . . et al. (1995). Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature, 376*(6538), 348-351. doi:10.1038/376348a0

Petrou, S., Ugur, M., Drummond, R. M., Singer, J. J., & Walsh, J. V., Jr. (1997). P2X7 purinoceptor expression in Xenopus oocytes is not sufficient to produce a pore-forming P2Z-like phenotype. *FEBS Lett, 411*(2-3), 339-345. doi:10.1016/s0014-5793(97)00700-x

Pini, G., Bigoni, S., Engerstrom, I. W., Calabrese, O., Felloni, B., Scusa, M. F., . . . group, E. (2012). Variant of Rett syndrome and CDKL5 gene: clinical and autonomic description of 10 cases. *Neuropediatrics, 43*(1), 37-43. doi:10.1055/s-0032-1308856

Plagnol, V., Curtis, J., Epstein, M., Mok, K. Y., Stebbings, E., Grigoriadou, S., . . . Nejentsev, S. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics, 28*(21), 2747-2754. doi:10.1093/bioinformatics/bts526

Portela, A., & Esteller, M. (2010). Epigenetic modifications and human disease. *Nat Biotechnol, 28*(10), 1057-1068. doi:10.1038/nbt.1685

Psoni, S., Sofocleous, C., Traeger-Synodinos, J., Kitsiou-Tzeli, S., Kanavakis, E., & Fryssira-Kanioura, H. (2010). Phenotypic and genotypic variability in four males with MECP2 gene sequence aberrations including a novel deletion. *Pediatr Res, 67*(5), 551-556. doi:10.1203/PDR.0b013e3181d4ecf7

Qu, D., Li, Q., Lim, H. Y., Cheung, N. S., Li, R., Wang, J. H., & Qi, R. Z. (2002). The protein SET binds the neuronal Cdk5 activator p35nck5a and modulates Cdk5/p35nck5a activity. *J Biol Chem, 277*(9), 7324-7332. doi:10.1074/jbc.M107270200

Raices, M., & D'Angelo, M. A. (2012). Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat Rev Mol Cell Biol, 13*(11), 687-699. doi:10.1038/nrm3461

Ramkumar, A., Jong, B. Y., & Ori-McKenney, K. M. (2018). ReMAPping the microtubule landscape: How phosphorylation dictates the activities of microtubule-associated proteins. *Dev Dyn, 247*(1), 138-155. doi:10.1002/dvdy.24599

Reichow, B., George-Puskar, A., Lutz, T., Smith, I. C., & Volkmar, F. R. (2015). Brief report: systematic review of Rett syndrome in males. *J Autism Dev Disord, 45*(10), 3377-3383. doi:10.1007/s10803-015-2519-1

Reiss, A. L., Faruque, F., Naidu, S., Abrams, M., Beaty, T., Bryan, R. N., & Moser, H. (1993). Neuroanatomy of Rett syndrome: a volumetric imaging study. *Ann Neurol, 34*(2), 227-234. doi:10.1002/ana.410340220

Rett, A. (1966). [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien Med Wochenschr, 116*(37), 723-726.

Revenkova, E., Focarelli, M. L., Susani, L., Paulis, M., Bassi, M. T., Mannini, L., . . . Musio, A. (2009). Cornelia de Lange syndrome mutations in SMC1A or SMC3 affect binding to DNA. *Hum Mol Genet, 18*(3), 418-427. doi:10.1093/hmg/ddn369

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Richardson, R., Splitt, M., Newbury-Ecob, R., Hulbert, A., Kennedy, J., Weber, A., & Study, D. D. D. (2018). SET de novo frameshift variants associated with developmental delay

and intellectual disabilities. *Eur J Hum Genet, 26*(9), 1306-1311. doi:10.1038/s41431-018-0199-y

Riviere, J. B., Ramalingam, S., Lavastre, V., Shekarabi, M., Holbert, S., Lafontaine, J., . . . Rouleau, G. A. (2011). KIF1A, an axonal transporter of synaptic vesicles, is mutated in hereditary sensory and autonomic neuropathy type 2. *Am J Hum Genet, 89*(2), 219-230. doi:10.1016/j.ajhg.2011.06.013

RK, C. Y., Merico, D., Bookman, M., J, L. H., Thiruvahindrapuram, B., Patel, R. V., . . . Scherer, S. W. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci, 20*(4), 602-611. doi:10.1038/nn.4524

Roelfsema, J. H., & Peters, D. J. (2007). Rubinstein-Taybi syndrome: clinical and molecular overview. *Expert Rev Mol Med, 9*(23), 1-16. doi:10.1017/S1462399407000415

Roelfsema, J. H., White, S. J., Ariyurek, Y., Bartholdi, D., Niedrist, D., Papadia, F., . . . Peters, D. J. (2005). Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet, 76*(4), 572-580. doi:10.1086/429130

Roux, J. C., Zala, D., Panayotis, N., Borges-Correia, A., Saudou, F., & Villard, L. (2012). Modification of Mecp2 dosage alters axonal transport through the Huntingtin/Hap1 pathway. *Neurobiol Dis, 45*(2), 786-795. doi:10.1016/j.nbd.2011.11.002

Sadedin, S. P., Dashnow, H., James, P. A., Bahlo, M., Bauer, D. C., Lonie, A., . . . Thorne, N. P. (2015). Cpipe: a shared variant detection pipeline designed for diagnostic settings. *Genome Med, 7*(1), 68. doi:10.1186/s13073-015-0191-x

Sadedin, S. P., Ellis, J. A., Masters, S. L., & Oshlack, A. (2018). Ximmer: a system for improving accuracy and consistency of CNV calling from exome data. *Gigascience, 7*(10). doi:10.1093/gigascience/giy112

Saitsu, H., Tohyama, J., Walsh, T., Kato, M., Kobayashi, Y., Lee, M., . . . Matsumoto, N. (2014). A girl with West syndrome and autistic features harboring a de novo TBL1XR1 mutation. *J Hum Genet, 59*(10), 581-583. doi:10.1038/jhg.2014.71

Sajan, S. A., Jhangiani, S. N., Muzny, D. M., Gibbs, R. A., Lupski, J. R., Glaze, D. G., . . . Neul, J. L. (2017). Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet Med, 19*(1), 13-19. doi:10.1038/gim.2016.42

Salpietro, V., Malintan, N. T., Llano-Rivas, I., Spaeth, C. G., Efthymiou, S., Striano, P., . . . Houlden, H. (2019). Mutations in the Neuronal Vesicular SNARE VAMP2 Affect Synaptic Membrane Fusion and Impair Human Neurodevelopment. *Am J Hum Genet, 104*(4), 721-730. doi:10.1016/j.ajhg.2019.02.016

Sandestig, A., Engström, K., Pepler, A., Danielsson, I., Odelberg-Johnsson, P., Biskup, S., . . . Stefanova, M. (2019). NUP188 Biallelic Loss of Function May Underlie a New Syndrome: Nucleoporin 188 Insufficiency Syndrome? *Molecular Syndromology*. doi:10.1159/000504818

Sanfeliu, A., Kaufmann, W. E., Gill, M., Guasoni, P., & Tropea, D. (2019). Transcriptomic Studies in Mouse Models of Rett Syndrome: A Review. *Neuroscience, 413*, 183-205. doi:10.1016/j.neuroscience.2019.06.013

Sawada, A., Takihara, Y., Kim, J. Y., Matsuda-Hashii, Y., Tokimasa, S., Fujisaki, H., . . . Hara, J. (2003). A congenital mutation of the novel gene LRRC8 causes agammaglobulinemia in humans. *J Clin Invest, 112*(11), 1707-1713. doi:10.1172/JCI18937

Sbalzarini, I. F., & Koumoutsakos, P. (2005). Feature point tracking and trajectory analysis for video imaging in cell biology. *J Struct Biol, 151*(2), 182-195. doi:10.1016/j.jsb.2005.06.002

Scala, E., Ariani, F., Mari, F., Caselli, R., Pescucci, C., Longo, I., . . . Renieri, A. (2005). CDKL5/STK9 is mutated in Rett syndrome variant with infantile spasms. *J Med Genet, 42*(2), 103-107. doi:10.1136/jmg.2004.026237

Scheper, G. C., van der Knaap, M. S., & Proud, C. G. (2007). Translation matters: protein synthesis defects in inherited disease. *Nat Rev Genet, 8*(9), 711-723. doi:10.1038/nrg2142

Schoberleitner, I., Mutti, A., Sah, A., Wille, A., Gimeno-Valiente, F., Piatti, P., . . . Lusser, A. (2019). Role for Chromatin Remodeling Factor Chd1 in Learning and Memory. *Front Mol Neurosci, 12*, 3. doi:10.3389/fnmol.2019.00003

Schonewolf-Greulich, B., Bisgaard, A. M., Duno, M., Jespersgaard, C., Rokkjaer, M., Hansen, L. K., . . . Tumer, Z. (2019). Mosaic MECP2 variants in males with classical Rett syndrome features, including stereotypical hand movements. *Clin Genet, 95*(3), 403-408. doi:10.1111/cge.13473

Schonewolf-Greulich, B., Bisgaard, A. M., Moller, R. S., Duno, M., Brondum-Nielsen, K., Kaur, S., . . . Tumer, Z. (2019). Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin Genet, 95*(2), 221-230. doi:10.1111/cge.13153

Schrödinger. (2010). The PyMOL Molecular Graphics System.

Schule, B., Armstrong, D. D., Vogel, H., Oviedo, A., & Francke, U. (2008). Severe congenital encephalopathy caused by MECP2 null mutations in males: central hypoxia and reduced neuronal dendritic structure. *Clin Genet, 74*(2), 116-126. doi:10.1111/j.1399-0004.2008.01005.x

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods, 11*(4), 361-362. doi:10.1038/nmeth.2890

Schwenk, J., Metz, M., Zolles, G., Turecek, R., Fritzius, T., Bildl, W., . . . Bettler, B. (2010). Native GABA(B) receptors are heteromultimers with a family of auxiliary subunits. *Nature, 465*(7295), 231-235. doi:10.1038/nature08964

Seddik, R., Jungblut, S. P., Silander, O. K., Rajalu, M., Fritzius, T., Besseyrias, V., . . . Bettler, B. (2012). Opposite effects of KCTD subunit domains on GABA(B) receptor-mediated desensitization. *J Biol Chem, 287*(47), 39869-39877. doi:10.1074/jbc.M112.412767

Seto, E., & Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb Perspect Biol, 6*(4), a018713. doi:10.1101/cshperspect.a018713

Sgado, P., Dunleavy, M., Genovesi, S., Provenzano, G., & Bozzi, Y. (2011). The role of GABAergic system in neurodevelopmental disorders: a focus on autism and epilepsy. *Int J Physiol Pathophysiol Pharmacol, 3*(3), 223-235.

Shahbazian, M. D., Antalffy, B., Armstrong, D. L., & Zoghbi, H. Y. (2002). Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. *Hum Mol Genet, 11*(2), 115-124.

Shahen, V. A., Cantrill, L. C., Sangani, N. B., Christodoulou, J., & Gold, W. A. (2018). A simple and efficient toolset for analysing mitochondrial trafficking in neuronal cells. *Acta Histochem, 120*(8), 797-805. doi:10.1016/j.acthis.2018.09.001

Sheikh, B. N. (2014). Crafting the brain - role of histone acetyltransferases in neural development and disease. *Cell Tissue Res, 356*(3), 553-573. doi:10.1007/s00441-014-1835-7

Shin, H., Wyszynski, M., Huh, K. H., Valtschanoff, J. G., Lee, J. R., Ko, J., . . . Kim, E. (2003). Association of the kinesin motor KIF1A with the multimodular protein liprin-alpha. *J Biol Chem, 278*(13), 11393-11401. doi:10.1074/jbc.M211874200

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res, 40*(Web Server issue), W452-457. doi:10.1093/nar/gks539

Simpson, M. A., Deshpande, C., Dafou, D., Vissers, L. E., Woollard, W. J., Holder, S. E., . . . Trembath, R. C. (2012). De novo mutations of the gene encoding the histone acetyltransferase KAT6B cause Genitopatellar syndrome. *Am J Hum Genet, 90*(2), 290-294. doi:10.1016/j.ajhg.2011.11.024

Skene, P. J., Illingworth, R. S., Webb, S., Kerr, A. R., James, K. D., Turner, D. J., . . . Bird, A. P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol Cell, 37*(4), 457-468. doi:10.1016/j.molcel.2010.01.030

Sokpor, G., Castro-Hernandez, R., Rosenbusch, J., Staiger, J. F., & Tuoc, T. (2018). ATP-Dependent Chromatin Remodeling During Cortical Neurogenesis. *Front Neurosci, 12*, 226. doi:10.3389/fnins.2018.00226

Spagnoli, C., Rizzi, S., Salerno, G. G., Frattini, D., & Fusco, C. (2019). Long-term follow-up until early adulthood in autosomal dominant, complex SPG30 with a novel KIF1A variant: a case report. *Ital J Pediatr, 45*(1), 155. doi:10.1186/s13052-019-0752-5

Stanhope, K., & Ross, J. (2015). Microtubules, MAPs, and Motor Proteins. In J. Ross & W. Marshall (Eds.), *Methods in Cell Biology* (Vol. 128, pp. 23-37).

Stembalska, A., Gil, J., & Pesz, K. A. (2011). [Rett syndrome. Classical form and preserved speech variant as a different phenotype effect of deletion with the same starting point in MeCP2 gene - report of 2 cases]. *Med Wieku Rozwoj, 15*(4), 445-450.

Stevens, C. A., & Krantz, I. (2018). 38th Annual David W. Smith Workshop on Malformations and Morphogenesis: Abstracts of the 2017 Annual Meeting. *Am J Med Genet A*, 1463-1536. doi:10.1002/ajmg.a.38698

Stevens, S. J. C., van der Schoot, V., Leduc, M. S., Rinne, T., Lalani, S. R., Weiss, M. M., . . . Brunner, H. G. (2018). De novo mutations in the SET nuclear proto-oncogene,

encoding a component of the inhibitor of histone acetyltransferases (INHAT) complex in patients with nonsyndromic intellectual disability. *Hum Mutat, 39*(7), 1014-1023. doi:10.1002/humu.23541

Stolerman, E. S., Smith, B., Chaubey, A., & Jones, J. R. (2016). CHD8 intragenic deletion associated with autism spectrum disorder. *Eur J Med Genet, 59*(4), 189-194. doi:10.1016/j.ejmg.2016.02.010

Subramaniam, B., Naidu, S., & Reiss, A. L. (1997). Neuroanatomy in Rett syndrome: cerebral cortex and posterior fossa. *Neurology, 48*(2), 399-407. doi:10.1212/wnl.48.2.399

Swiech, L., Blazejczyk, M., Urbanska, M., Pietruszka, P., Dortland, B. R., Malik, A. R., . . . Jaworski, J. (2011). CLIP-170 and IQGAP1 cooperatively regulate dendrite morphology. *J Neurosci, 31*(12), 4555-4568. doi:10.1523/JNEUROSCI.6582-10.2011

Tao, J., Hu, K., Chang, Q., Wu, H., Sherman, N. E., Martinowich, K., . . . Sun, Y. E. (2009). Phosphorylation of MeCP2 at Serine 80 regulates its chromatin association and neurological function. *Proc Natl Acad Sci U S A, 106*(12), 4882-4887. doi:10.1073/pnas.0811648106

ten Klooster, J. P., Leeuwen, I., Scheres, N., Anthony, E. C., & Hordijk, P. L. (2007). Rac1-induced cell migration requires membrane recruitment of the nuclear oncogene SET. *EMBO J, 26*(2), 336-345. doi:10.1038/sj.emboj.7601518

Teng, X., Aouacheria, A., Lionnard, L., Metz, K. A., Soane, L., Kamiya, A., & Hardwick, J. M. (2019). KCTD: A new gene family involved in neurodevelopmental and neuropsychiatric disorders. *CNS Neurosci Ther, 25*(7), 887-902. doi:10.1111/cns.13156

Theerthagiri, G., Eisenhardt, N., Schwarz, H., & Antonin, W. (2010). The nucleoporin Nup188 controls passage of membrane proteins across the nuclear pore complex. *J Cell Biol, 189*(7), 1129-1142. doi:10.1083/jcb.200912045

Topcu, M., Akyerli, C., Sayi, A., Toruner, G. A., Kocoglu, S. R., Cimbis, M., & Ozcelik, T. (2002). Somatic mosaicism for a MECP2 mutation associated with classic Rett syndrome in a boy. *Eur J Hum Genet, 10*(1), 77-81. doi:10.1038/sj.ejhg.5200745

Toriyama, M., Shimada, T., Kim, K. B., Mitsuba, M., Nomura, E., Katsuta, K., . . . Inagaki, N. (2006). Shootin1: A protein involved in the organization of an asymmetric signal for neuronal polarization. *J Cell Biol, 175*(1), 147-157. doi:10.1083/jcb.200604160

Townend, G. S., Bartl-Pokorny, K. D., Sigafoos, J., Curfs, L. M., Bolte, S., Poustka, L., . . . Marschik, P. B. (2015). Comparing social reciprocity in preserved speech variant and typical Rett syndrome during the early years of life. *Res Dev Disabil, 43-44*, 80-86. doi:10.1016/j.ridd.2015.06.008

Tyagi, M., Imam, N., Verma, K., & Patel, A. K. (2016). Chromatin remodelers: We are the drivers!! *Nucleus, 7*(4), 388-404. doi:10.1080/19491034.2016.1211217

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res, 40*(15), e115. doi:10.1093/nar/gks596

Van Driesche, S. J., & Martin, K. C. (2018). New frontiers in RNA transport and local translation in neurons. *Dev Neurobiol, 78*(3), 331-339. doi:10.1002/dneu.22574

Veeramah, K. R., Johnstone, L., Karafet, T. M., Wolf, D., Sprissler, R., Salogiannis, J., . . . Hammer, M. F. (2013). Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia, 54*(7), 1270-1281. doi:10.1111/epi.12201

Vegas, N., Cavallin, M., Maillard, C., Boddaert, N., Toulouse, J., Schaefer, E., . . . Bahi-Buisson, N. (2018). Delineating FOXG1 syndrome: From congenital microcephaly to hyperkinetic encephalopathy. *Neurol Genet, 4*(6), e281. doi:10.1212/NXG.0000000000000281

Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L., & Vriend, G. (2010). Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics, 11*, 548. doi:10.1186/1471-2105-11-548

Vidal, S., Brandi, N., Pacheco, P., Maynou, J., Fernandez, G., Xiol, C., . . . Armstrong, J. (2019). The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome. *Eur J Paediatr Neurol, 23*(4), 609-620. doi:10.1016/j.ejpn.2019.04.006

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

Viosca, J., Lopez-Atalaya, J. P., Olivares, R., Eckner, R., & Barco, A. (2010). Syndromic features and mild cognitive impairment in mice with genetic reduction on p300 activity: Differential contribution of p300 and CBP to Rubinstein-Taybi syndrome etiology. *Neurobiol Dis, 37*(1), 186-194. doi:10.1016/j.nbd.2009.10.001

Vorobiov, D., Levin, G., Lotan, I., & Dascal, N. (1998). Agonist-independent inactivation and agonist-induced desensitization of the G protein-activated K+ channel (GIRK) in Xenopus oocytes. *Pflugers Arch, 436*(1), 56-68. doi:10.1007/s004240050604

Vuillaume, M. L., Jeanne, M., Xue, L., Blesson, S., Denomme-Pichon, A. S., Alirol, S., . . . Toutain, A. (2018). A novel mutation in the transmembrane 6 domain of GABBR2 leads to a Rett-like phenotype. *Ann Neurol, 83*(2), 437-439. doi:10.1002/ana.25155

Wang, D., Kon, N., Lasso, G., Jiang, L., Leng, W., Zhu, W. G., . . . Gu, W. (2016). Acetylation-regulated interaction between p53 and SET reveals a widespread regulatory mode. *Nature, 538*(7623), 118-122. doi:10.1038/nature19759

Wang, E. T., Taliaferro, J. M., Lee, J. A., Sudhakaran, I. P., Rossoll, W., Gross, C., . . . Bassell, G. J. (2016). Dysregulation of mRNA Localization and Translation in Genetic Disease. *J Neurosci, 36*(45), 11418-11426. doi:10.1523/JNEUROSCI.2352-16.2016

Wang, J., Weaver, I. C., Gauthier-Fisher, A., Wang, H., He, L., Yeomans, J., . . . Miller, F. D. (2010). CBP histone acetyltransferase activity regulates embryonic neural differentiation in the normal and Rubinstein-Taybi syndrome brain. *Dev Cell, 18*(1), 114-125. doi:10.1016/j.devcel.2009.10.023

Wang, J., Xie, R., Kou, X., Liu, Y., Qi, C., Liu, R., . . . Gao, X. (2019). A protein phosphatase 2A deficit in the hippocampal CA1 area impairs memory extinction. *Mol Brain, 12*(1), 51. doi:10.1186/s13041-019-0469-9

Wang, J., Zhang, Q., Chen, Y., Yu, S., Wu, X., & Bao, X. (2019). Rett and Rett-like syndrome: Expanding the genetic spectrum to KIF1A and GRIN1 gene. *Mol Genet Genomic Med*, e968. doi:10.1002/mgg3.968

Wapenaar, H., & Dekker, F. J. (2016). Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin Epigenetics, 8*, 59. doi:10.1186/s13148-016-0225-2

Wickramasinghe, V. O., McMurtrie, P. I., Mills, A. D., Takei, Y., Penrhyn-Lowe, S., Amagase, Y., . . . Laskey, R. A. (2010). mRNA export from mammalian cell nuclei is dependent on GANP. *Curr Biol, 20*(1), 25-31. doi:10.1016/j.cub.2009.10.078

Willemsen, M. H., Ba, W., Wissink-Lindhout, W. M., de Brouwer, A. P., Haas, S. A., Bienek, M., . . . Kleefstra, T. (2014). Involvement of the kinesin family members KIF4A and KIF5C in intellectual disability and synaptic function. *J Med Genet, 51*(7), 487-494. doi:10.1136/jmedgenet-2013-102182

Williams, E. C., Zhong, X., Mohamed, A., Li, R., Liu, Y., Dong, Q., . . . Chang, Q. (2014). Mutant astrocytes differentiated from Rett syndrome patients-specific iPSCs have adverse effects on wild-type neurons. *Hum Mol Genet, 23*(11), 2968-2980. doi:10.1093/hmg/ddu008

Williamson, S. L., & Christodoulou, J. (2006). Rett syndrome: new clinical and molecular insights. *Eur J Hum Genet, 14*(8), 896-903. doi:10.1038/sj.ejhg.5201580

Wong, L. C., Singh, S., Wang, H. P., Hsu, C. J., Hu, S. C., & Lee, W. T. (2019). FOXG1-Related Syndrome: From Clinical to Molecular Genetics and Pathogenic Mechanisms. *Int J Mol Sci, 20*(17). doi:10.3390/ijms20174176

Xu, X., He, C., Zhang, Z., & Chen, Y. (2006). MKLP1 requires specific domains for its dendritic targeting. *J Cell Sci, 119*(Pt 3), 452-458. doi:10.1242/jcs.02750

Xu, X., Kozikowski, A. P., & Pozzo-Miller, L. (2014). A selective histone deacetylase-6 inhibitor improves BDNF trafficking in hippocampal neurons from Mecp2 knockout mice: implications for Rett syndrome. *Front Cell Neurosci, 8*, 68. doi:10.3389/fncel.2014.00068

Yasuhara, N., Shibazaki, N., Tanaka, S., Nagai, M., Kamikawa, Y., Oe, S., . . . Yoneda, Y. (2007). Triggering neural differentiation of ES cells by subtype switching of importin-alpha. *Nat Cell Biol, 9*(1), 72-79. doi:10.1038/ncb1521

Yi, F., Danko, T., Botelho, S. C., Patzke, C., Pak, C., Wernig, M., & Sudhof, T. C. (2016). Autism-associated SHANK3 haploinsufficiency causes Ih channelopathy in human neurons. *Science, 352*(6286), aaf2669. doi:10.1126/science.aaf2669

Ylikallio, E., Kim, D., Isohanni, P., Auranen, M., Kim, E., Lonnqvist, T., & Tyynismaa, H. (2015). Dominant transmission of de novo KIF1A motor domain variant underlying pure spastic paraplegia. *Eur J Hum Genet, 23*(10), 1427-1430. doi:10.1038/ejhg.2014.297

Yoo, Y., Jung, J., Lee, Y. N., Lee, Y., Cho, H., Na, E., . . . Choi, M. (2017). GABBR2 mutations determine phenotype in rett syndrome and epileptic encephalopathy. *Ann Neurol, 82*(3), 466-478. doi:10.1002/ana.25032

Yoshikawa, K., Kuwahara, M., Saigoh, K., Ishiura, H., Yamagishi, Y., Hamano, Y., . . . Kusunoki, S. (2019). The novel de novo mutation of KIF1A gene as the cause for Spastic paraplegia 30 in a Japanese case. *eNeurologicalSci, 14*, 34-37. doi:10.1016/j.ensci.2018.11.026

Yu, W., Cook, C., Sauter, C., Kuriyama, R., Kaplan, P. L., & Baas, P. W. (2000). Depletion of a microtubule-associated motor protein induces the loss of dendritic identity. *J Neurosci, 20*(15), 5782-5791.

Yu, W., Sharp, D. J., Kuriyama, R., Mallik, P., & Baas, P. W. (1997). Inhibition of a mitotic motor compromises the formation of dendrite-like processes from neuroblastoma cells. *J Cell Biol, 136*(3), 659-668.

Yue, Y., Blasius, T. L., Zhang, S., Jariwala, S., Walker, B., Grant, B. J., . . . Verhey, K. J. (2018). Altered chemomechanical coupling causes impaired motility of the kinesin-4 motors KIF27 and KIF7. *J Cell Biol, 217*(4), 1319-1334. doi:10.1083/jcb.201708179

Zachariah, R. M., & Rastegar, M. (2012). Linking epigenetics to human disease and Rett syndrome: the emerging novel and challenging concepts in MeCP2 research. *Neural Plast, 2012*, 415825. doi:10.1155/2012/415825

Zappella, M., Meloni, I., Longo, I., Hayek, G., & Renieri, A. (2001). Preserved speech variants of the Rett syndrome: molecular and clinical analysis. *Am J Med Genet, 104*(1), 14-22.

Zaret, K. S., & Mango, S. E. (2016). Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev, 37*, 76-81. doi:10.1016/j.gde.2015.12.003

Zhang, Q., Wang, J., Li, J., Bao, X., Zhao, Y., Zhang, X., . . . Wu, X. (2017). Novel FOXG1 mutations in Chinese patients with Rett syndrome or Rett-like mental retardation. *BMC Med Genet, 18*(1), 96. doi:10.1186/s12881-017-0455-y

Zheng, S., Abreu, N., Levitz, J., & Kruse, A. C. (2019). Structural basis for KCTD-mediated rapid desensitization of GABAB signalling. *Nature, 567*(7746), 127-131. doi:10.1038/s41586-019-0990-0

Zhou, Z., Hong, E. J., Cohen, S., Zhao, W. N., Ho, H. Y., Schmidt, L., . . . Greenberg, M. E. (2006). Brain-specific phosphorylation of MeCP2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron, 52*(2), 255-269. doi:10.1016/j.neuron.2006.09.037

Zhu, Y. C., & Xiong, Z. Q. (2019). Molecular and Synaptic Bases of CDKL5 Disorder. *Dev Neurobiol, 79*(1), 8-19. doi:10.1002/dneu.22639

Zocchi, L., & Sassone-Corsi, P. (2012). SIRT1-mediated deacetylation of MeCP2 contributes to BDNF expression. *Epigenetics, 7*(7), 695-700. doi:10.4161/epi.20733

Zoghbi, H. Y. (2003). Postnatal neurodevelopmental disorders: meeting at the synapse? *Science, 302*(5646), 826-830. doi:10.1126/science.1089071

Appendices

Appendices

Appendix A

Genes, Primers and on-line tools

List A1: Known RTT and RTT like genes

ACTL6B, ADSL, ANXA11, ASXL3, ATP6V0A1, ATRX, BRAF, CACNA1D, CACNA1G, ***CACNA1I****, CAMK2B,* ***CDKL5****, CHD4,* ***CHRNA5****, CLTC, COL4A1, CTNNB1, CUX2, DEPDC5, DMXL2, DNMT3A, EEF1A2, EHMT1, EIF2B2, EIF4G1, FAM151A,* ***FOXG1****,* ***GABBR2****, GABRA1, GABRB2, GABRD, GABRG2, GRIA3, GRIN1, GRIN2A,* ***GRIN2B****,* ***HCN1****, HDAC1, HDAC4, HDAC8, HECW2 , HERC2, HTT, IMPDH2,* ***IQSEC2****, IQGAP3, IRF2BPL, ITPR1, JMJD1C, KANSL1, KCNA2,* ***KCNB1****, KCNJ10,* ***KCNQ2****, KIF1A, KIF4B, LAMB2, MAST1, MAST3, MBD5,* ***MECP2****,* ***MEF2C****, NALCN,* ***NCOR2****, NR2F1, NTNG1, OSBP, PDHA1, PPT1, RHOBTB2. RRN3, SATB2,* ***SCN1A****,* ***SCN2A****,* ***SCN8A****, SLC6A1, SDHA.* ***SHANK3****, SHROOM4, SLC2A1, SLC35A2, SLC39A13, SLC6A1, SLC9A6, SMARCA1,* ***SMC1A****, ST3GAL5,* ***STXBP1****,* ***SYNGAP1****, SYT1, TBL1XR1,* ***TCF4****, TRRAP,* ***UBE3A****, UBTF, USP8,* ***VAMP2****,* ***WDR45****, ZEB2, ZNF238.*

Note: Bold: ≥2 individuals; Bold and underlines: ≥4 individuals reported in the literature.

Gene list created from (Cogliati et al., 2019; Henriksen et al., 2018; Salpietro et al., 2019; Schonewolf-Greulich, Bisgaard, Moller, et al., 2019; Vidal, Brandi, et al., 2019; Vidal, Xiol, et al., 2019; Vuillaume et al., 2018)

List A2: List of 136 genes on custom designed Agilent SureSelect Panel

ACTL6B, ADAM23, ARX, ATRX, AUTS2, BRAF, BRD1, CDKL5, CELSR1, CHD4, CHL1, CNTN6, CNTNAP2, CREB1, CTNNB1, EEF1A2, EHMT1, EIF2B2, FAM151A, FAT3, FOLR1, FOXG1, GABRB2, GABRD, GNB1, GRIN2A, GRIN2B, HDAC1, HERC2, IMPDH2, IQGAP3, IQSEC2, KCNA2, KCNJ10, KCNQ2, KIF1A, KIF1B, KIF1C, KIF2A, KIF2B, KIF2C, KIF3A, KIF3B, KIF3C, KIF4A, KIF4B, KIF5A, KIF5B, KIF5C, KIF6, KIF7, KIF9, KIF11, KIF12, KIF13A, KIF13B, KIF14, KIF15, KIF16A, KIF16B, KIF17, KIF18A, KIF18B, KIF19, KIF20A, KIF20B, KIF21A, KIF21B, KIF22, KIF23, KIF24, KIF25, KIF26A, KIF26B, KIF27, KIFC1, KIFC2, KIFC3, KIFAP3, KLF7, LAMB2, LANCL1, LRRC40, MAP2, MBD5, MECP2, MEF2C, NAGA, NCOR2, NDNL2, NRP2, NRXN1, NTNG1, OPHN1, OTUD7A, PCDH19, PQBP1, PLCB1, PLXNB2, PNKP, PPP6R2, PRKD1, PTPN4, SATB2, SBF1, SCN8A, SCO2, SCUBE1, SHANK3, SHROOM4, SLC2A1, SLC9A6, SLC19A3,

SLC25A22, SLC35A2, SLC39A13, SLC6A1, SMARCB1, SMC1A, SPECC1L, SPTAN1, STXBP1, TBL1XR1, TCF20, TCF4, TRPM1, TRRAP, TUBGCP6, WDR45, WNT7B, UBE3A, UNC80, ZEB2, ZFX, ZNF238, ZNF536

List A3: Victorian Clinical Genetics Services pre-curated intellectual disability gene-list

Victorian Clinical Genetics Services is a NATA-accredited genetic diagnosis lab that has created phenotype specific gene lists in order to improve the diagnostic yield. The following list of genes can be found at:

https://www.vcgs.org.au/sites/default/files/downloads/TGW024_genelist_V3.pdf

AARS, AASS, ABCC9, ABCD1, ABCD4, ABHD5, ACAD9, ACADM, ACADS, ACADSB, ACAT1, ACO2, ACOX1, ACSL4, ACTB, ACTG1, ACTL6A, ACY1, ADAR, ADCK3, ADK, ADNP, ADSL, AFF2, AFF3, AFF4, AGL, AHDC1, AHI1, AIFM1, AIMP1, AKAP6, AKT3, ALDH18A1, ALDH3A2, ALDH4A1, ALDH5A1, ALDH7A1, ALG1, ALG11, ALG12, ALG13, ALG2, ALG3, ALG6, ALG8, ALG9, ALMS1, ALX4, AMER1, AMPD2, AMT, ANK3, ANKRD11, AP1S1, AP1S2, AP3B1, AP3B2, AP4B1, AP4E1, AP4M1, AP4S1, APOPT1, AR, ARCN1, ARFGEF2, ARG1, ARHGEF6, ARHGEF9, ARID1A, ARID1B, ARID2, ARL13B, ARL6, ARMC9, ARSA, ARSB, ARSE, ARV1, ARX, ASAH1, ASH1L, ASMT, ASPA, ASPM, ASS1, ASXL1, ASXL2, ASXL3, ATAD3A, ATIC, ATL1, ATM, ATP13A2, ATP1A2, ATP6AP2, ATP6V0A2, ATP6V1B2, ATP7A, ATR, ATRX, AUH, AUTS2, B3GALNT2, B3GALTL, B4GALNT1, B4GALT7, BBS1, BBS10, BBS12, BBS2, BBS4, BBS5, BBS7, BBS9, BCAP31, BCKDHA, BCKDHB, BCKDK, BCL11A, BCOR, BCS1L, BIN1, BLM, BMP4, BOLA3, BPTF, BRAF, BRF1, BRIP1, BRPF1, BRWD3, BSCL2, BTD, BUB1BC12orf4, C12orf57, C12orf65, C2CD3, C3orf58, C5orf42, CA5A, CACNA1A, CACNA1C, CACNA1D, CACNG2, CAMK2A, CAMK2B, CAMTA1, CANT1, CASC5, CASK, CBL, CBS, CC2D1A, CC2D2A, CCBE1, CCDC22, CCDC41, CCDC88A, CCDC88C, CCND2, CDC42, CDC6, CDH11, CDH15, CDK10, CDK13, CDK5R1, CDK5RAP2, CDKL5, CDKN1C, CDON, CENPF, CENPJ, CEP135, CEP152, CEP290, CEP41, CEP57, CHAMP1, CHD2, CHD4, CHD7, CHD8, CHKB, CHMP1A, CHRNA4, CIC, CIT, CKAP2L, CLCN4, CLIC2, CLN3, CLN5, CLN6, CLN8, CLP1, CLPB, CLPP, CLTC, CNKSR2, CNNM2, CNOT3, CNTN3, CNTNAP2, COA3, COASY, COG1, COG4, COG5, COG7, COG8, COL1A2, COL4A1, COL4A2, COL4A3BP, COLEC10, COLEC11, COQ4, COQ5, COX10, COX15, COX6B1, CPS1, CRADD, CRB2, CRBN, CREBBP, CSNK2A1, CSNK2B, CSPP1, CSTB, CTC1, CTCF, CTDP1, CTNNB1, CTNND2, CTSA, CTSD, CUL4B, CWC27, CYB5R3, CYC1, CYP27A1, D2HGDH, DAG1, DARS, DARS2, DBT, DCAF17, DCHS1, DCX, DDC, DDHD2, DDOST, DDX11, DDX3X, DEAF1, DEPDC5, DHCR24, DHCR7, DHFR, DHTKD1, DHX30, DIAPH1, DIP2B, DIS3L2, DKC1, DLD, DLG3, DLGAP2, DMD, DNAJC12, DNAJC19, DNAJC3, DNM1, DNMT3A, DNMT3B, DOCK3, DOCK7, DOCK8, DOLK, DPAGT1, DPF2, DPM1, DPM2, DPM3, DPP6, DPYD, DSCAM, DYNC1H1, DYRK1A, EBF3, EBP, EED, EEF1A2, EFNB1, EFTUD2, EHMT1, EIF2AK3, EIF2S3, EIF4A3, ELAC2, ELOVL4, ELP2, EML1, EMX2, EP300, EPB41L1, EPG5, EPT1, ERCC1, ERCC2, ERCC3, ERCC5, ERCC6, ERCC6L2, ERCC8, ESCO2, ETHE1, EXOSC3, EXTL3, EZH2, FAM111A, FAM126A, FAM20C, FANCB, FANCG, FAR1, FAT4, FBXL4, FDXR, FGD1, FGF12, FGF14, FGFR1, FGFR2, FGFR3, FH, FIBP, FIG4, FKRP, FKTN, FLNA, FLVCR2, FMN2, FMR1, FOLR1, FOXG1, FOXP1, FOXP2, FOXRED1, FTCD, FTO, FTSJ1, FUCA1, FZD3, G6PC3, GABRA1, GABRB2, GABRB3, GABRG2, GALC, GALE, GAMT, GAN, GATAD2B, GATM, GBA, GCDH, GCH1, GDI1, GEMIN4, GFAP, GFER, GFM1, GJC2, GK, GLB1,

GLDC, GLI2, GLI3, GLRA1, GLUL, GLYCTK, GM2A, GMPPA, GMPPB, GNAI1, GNAO1, GNAS, GNB1, GNPAT, GNPTAB, GNPTG, GNS, GPAA1, GPC3, GPR56, GPT2, GRIA3, GRIA4, GRID2, GRIK2, GRIN1, GRIN2A, GRIN2B, GRM1, GSS, GTDC2, GTF2H5, GTF3C3, GTPBP2, GTPBP3, GUSB, H3F3A, H3F3B, HACE1, HADHA, HARS2, HAX1, HCCS, HCFC1, HCN1, HDAC4, HDAC8, HECW2, HEPACAM, HERC1, HERC2, HESX1, HEXA, HEXB, HGSNAT, HIBCH, HIST1H1E, HIST1H4C, HIVEP2, HLCS, HMGCL, HNRNPH2, HNRNPK, HNRNPU, HOXA1, HPD, HPRT1, HRAS, HSD17B10, HSD17B4, HSPD1, HTRA2, HUWE1, IARS, IDH2, IDS, IDUA, IER3IP1, IFIH1, IFT172, IGBP1, IGF1, IGF1R, IKBKG, IL1RAPL1, INPP5E, INPP5K, INSR, INTS1, INTS8, IQSEC2, IRF2BPL, IRX5, ISCA2, ISPD, ITGA7, ITPA, ITPR1, IVD, JAM3, KANSL1, KAT5, KAT6A, KAT6B, KATNAL2, KCNA2, KCNB1, KCNC1, KCNH1, KCNJ10, KCNJ11, KCNJ6, KCNK9, KCNQ2, KCNQ3, KCNQ5, KCNT1, KCTD3, KCTD7, KDM1A, KDM5B, KDM5C, KDM6A, KIAA0586, KIAA1109, KIAA1279, KIAA2022, KIDINS220, KIF11, KIF14, KIF1A, KIF4A, KIF5A, KIF5C, KIF7, KLHL7, KMT2A, KMT2C, KMT2D, KMT2E, KPTN, KRAS, L1CAM, L2HGDH, LAMA1, LAMA2, LAMB1, LAMC3, LAMP2, LARGE, LARP7, LARS2, LAS1L, LHX3, LIG4, LINS, LMBRD1, LNP1, LONP1, LRP2, LRP5, LRPPRC, MAB21L2, MADD, MAF, MAGEL2, MAN1B1, MAN2B1, MANBA, MAOA, MAP2K1, MAP2K2, MASP1, MAT1A, MBD5, MBOAT7, MBTPS2, MCCC1, MCCC2, MCOLN1, MCPH1, MDH2, MECP2, MED12, MED13L, MED17, MED23, MED25, MEF2C, MEIS2, MFF, MFSD8, MGAT2, MICU1, MID1, MKKS, MKS1, MLC1, MLYCD, MMAA, MMAB, MMACHC, MMADHC, MOCS1, MOCS2, MOGS, PDU1, MPI, MPLKIP, MRPS22, MSL3, MSMO1, MTFMT, MTHFR, MTOR, MTR, MTRR, MUT, MVK, MYCN, MYO5A, MYT1L, NAA10, NAA15, NACC1, NAGA, NAGLU, NALCN, NANS, NBN, NCKAP1, NDE1, NDP, NDST1, NDUFA1, NDUFAF5, NDUFS1, NDUFS4, NDUFS7, NDUFS8, NDUFV1, NEU1, NF1, NFIA, NFIX, NFU1, NGF, NGLY1, NHEJ1, NHP2, NHS, NIPBL, NKX21, NLGN3, NLGN4X, NONO, NPC1, NPC2, NPHP1, NPHP3, NR2F1, NRAS, NRXN1, NSD1, NSDHL, NSUN2, NT5C2, NTNG1, NTRK1, NUBPL, NUP188, OCLN, OCRL, OFD1, OGT, OPA3, OPHN1, ORC1, OTC, OTUD6B, OTX2, PACS1, PAFAH1B1, PAH, PAK3, PARN, PAX6, PAX7, PAX8, PBX1, PC, PCCA, PCCB, PCDH10, PCDH19, PCLO, PCNT, PDE4D, PDGFRB, PDHA1, PDHX, PDSS1, PDSS2, PEPD, PET100, PEX1, PEX10, PEX11B, PEX12, PEX13, PEX14, PEX16, PEX19, PEX2, PEX26, PEX3, PEX5, PEX6, PEX7, PGAP1, PGAP2, PGAP3, PGK1, PGM3, PHF6, PHF8, PHGDH, PIEZO2, PIGA, PIGC, PIGL, PIGN, PIGO, PIGT, PIGV, PIGW, PIGY, PIK3CA, PIK3R2, PLA2G6, PLAA, PLCB1, PLK4, PLP1, PMM2, PNKP, PNPLA6, POGZ, POLG, POLR3A, POLR3B, POMGNT1, POMT1, POMT2, PORCN, POU1F1, PPM1D, PPP1CB, PPP1R15B, PPP2R1A, PPP2R5D, PPP3CA, PPT1, PQBP1, PRKD1, PRMT7, PRODH, PROSC, PRPS1, PRRT2, PRSS12, PRUNE, PSAT1, PSMD12, PSPH, PTCH1, PTCHD1, PTDSS1, PTEN, PTF1A, PTPN11, PTPN23, PTS, PUF60, PURA, PUS1, PYCR1, PYCR2, QARS, QDPR, QRICH1, RAB11B, RAB18, RAB23, RAB39B, RAB3GAP1, RAB3GAP2, RAC1, RAD21, RAF1, RAI1, RARB, RARS2, RBBP8, RBFOX1, RBM10, RELN, RERE, RFT1, RHEB, RIMS1, RIT1, RLIM, RMND1, RNASEH2A, RNASEH2B, RNASEH2C, RNASET2, RNF125, RNF135, ROGDI, ROR2, RPGRIP1L, RPL10, RPS23, RPS6KA3, RRM2B, RTEL1, RTTN, SACS, SAMD9, SAMHD1, SATB2, SC5D, SCN1A, SCN2A, SCN3A, SCN8A, SCO1, SCO2, SDCCAG8, SDHA, SDHAF1, SERAC1, SET, SETBP1, SETD1B, SETD2, SETD5, SGPL1, SGSH, SHANK1, SHANK2, SHANK3, SHH, SHOC2, SHROOM4, SIK1, SIL1, SIN3A, SIX3, SKI, SLC12A5, SLC12A6, SLC13A5, SLC16A2, SLC17A5, SLC19A3, SLC25A1, SLC25A12, SLC25A15, SLC25A22, SLC25A24, SLC2A1, SLC2A2, SLC33A1, SLC35A1, SLC35A2, SLC35C1, SLC39A14, SLC39A8, SLC45A1, SLC46A1, SLC4A4, SLC5A5, SLC6A1, SLC6A17,

SLC6A19, SLC6A3, SLC6A4, SLC6A8, SLC6A9, SLC7A7, SLC9A6, SLC9A9, SLX4, SMAD4, SMARCA2, SMARCA4, SMARCB1, SMARCE1, SMC1A, SMC3, SMOC1, SMPD1, SMPD4, SMS, SNAP29, SNIP1, SNORD118, SNRPB, SNX14, SOBP, SON, SOS1, SOX10, SOX11, SOX2, SOX3, SOX5, SOX9, SPAST, SPATA5, SPECC1L, SPG11, SPR, SPRED1, SPRTN, SPTAN1, SPTBN2, SRCAP, SRD5A3, SRPX2, SSR4, ST3GAL3, STAG1, STAMBP, STIL, STRA6, STX11, STX1B, STXBP1, SUCLG1, SUOX, SURF1, SUV420H1, SYN1, SYNGAP1, SYNJ1, SYP, SYT14, SZT2"TAF1, TAF13, TAF2, TAF6, TANGO2, TAOK1, TAT, TAZ, TBC1D23, TBC1D24, TBC1D7, TBCD, TBCE, TBCK, TBL1XR1, TBR1, TBX1, TCF20, TCF4, TCN2, TCTN2, TECPR2, TECR, TGIF1, TH, THOC2, THOC6, THRA, THRB, TINF2, TLK2, TMCO1, TMEM165, TMEM216, TMEM237, TMEM240, TMEM260, TMEM5, TMEM67, TMEM70, TMLHE, TMTC3, TNIK, TNRC6B, TOE1, TPH2, TPK1TRAK1, TRAPPC12, TRAPPC6B, TRAPPC9, TREX1, TRIM32, TRIO, TRIP12, TRIP13, TRIT1, TRMT10A, TSC1, TSC2, TSEN2, TSEN34, TSEN54, TSFM, TSHB, TSHR, TSPAN7, TTC19, TTC37, TTC8, TTI2, TUBA1A, TUBA8, TUBB, TUBB2A, TUBB2B, TUBB3, TUBB4A, TUBG1, TUBGCP6, TUSC3, TWIST1, TXNL4A, UBA5, UBE2A, UBE3A, UBE3B, UBR1, UBR4, UBTF, UMPS, UNC13A, UNC80, UPB1, UPF3B, UROC1, USP18, USP9X, VLDLR, VPS13B, VPS53, VRK1, WAC, WDFY3, WDPCP, WDR26, WDR45, WDR45B, WDR62, WDR73, WDR81, WFS1, WWOX, XIST, XPNPEP3, XRCC4, YAP1, YWHAG, YY1, ZBTB16, ZBTB18, ZBTB20, ZBTB24, ZC4H2, ZDHHC9, ZEB2, ZFYVE26, ZIC2, ZMYND11, ZNF711, ZNF81, ZSWIM6

Table A1: List of primers used in this thesis.

S. No.	Primer name	Sequence (5'-3')	Application	Fragment size	Tm	%GC
1	MECP2_JC001_F	CACGGAAGCTTAAGCAAAGG	Variant confirmation	227	61	55
2	MECP2_JC001_R	CTTCACGGCTTTCTTTTTGG	Variant confirmation		60	45
3	MECP2_BBZ003_F	CACGGAAGCTTAAGCAAAGG	Variant confirmation	227	61	55
4	MECP2_BBZ003_R	CTTCACGGCTTTCTTTTTGG	Variant confirmation		60	45
5	MECP2_BBZ006_F	CACGGAAGCTTAAGCAAAGG	Variant confirmation	227	61	55
6	MECP2_BBZ006_R	CTTCACGGCTTTCTTTTTGG	Variant confirmation		60	45
7	MECP2_JC006_F	GCTCTGCTGGGAAGTATGATG	Variant confirmation	568	60	52
8	MECP2_JC006_R	CTTCACGGCTTTCTTTTTGG	Variant confirmation		60	45
9	CHD8_JC002_F	CCCAGTAGTTGTCCAAGATAACATTC	Variant confirmation	675	61	42
10	CHD8_JC002_R	TCCCATTGGGTTCAGTCAAG	Variant confirmation		61	50
11	KIF23_JC002_F	TTTTTGCGTAACACATTTGG	Variant confirmation	273	57	35
12	KIF23_JC002_R	TTTCCCGACAAAAACTCAAG	Variant confirmation		58	40
13	KAT6A_BBZ001_F	TGTAAGACCTTGTAACCTGTCTTTG	Variant confirmation	193	59	40
14	KAT6A_BBZ001_R	ATTGGTTTGCGGCTCTTG	Variant confirmation		60	50
15	KCTD16_BBZ005_F	TTCTACCGTGAGCCTTCCAG	Variant confirmation	375	60	55
16	KCTD16_BBZ005_R	GACAGGGCCCTTCTTGATG	Variant confirmation		61	58
17	EEF1A2_JC005_F	GGCCACTCTGCTGTAACAAG	Variant confirmation	1151	59	55
18	EEF1A2_JC005_R	CACAGGAGACCCTACCATCC	Variant confirmation		59	60
19	NSD1_JC007_F	TTTATCCAGCCTTGGGTCAG	Variant confirmation	809	60	50
20	NSD1_JC007_R	TGGGCGTCTGTCTTGTACTG	Variant confirmation		60	55
21	ZNF557_JC007_F	TATGGCTGCAGTGAATGTGG	Variant confirmation	600	61	50
22	ZNF557_JC007_R	GAAGATTGACTGAGTGCGGAAC	Variant confirmation		61	50
23	CNTNAP2_JC007_F	CCACAGGGAAGATTTGGTTC	Variant confirmation	600	59	50
24	CNTNAP2_JC007_R	ATTTGTTGACGCTTTGCTTG	Variant confirmation		59	40
25	CNTNAP4_JC007_F	AAGTTCCATAGAGGAATACAGTCG	Variant confirmation	450	58	42
26	CNTNAP4_JC007_R	AAGAGATTGGTAAACAAGATGCAG	Variant confirmation		59	38
27	SCN9A_JC007_F	CGTATTTTGTGTCCCCTACCC	Variant confirmation	246	60	52
28	SCN9A_JC007_R	TGGCGTAGATGAACATGACC	Variant confirmation		60	50
29	SYT16_JC007_F	GGCCAATGGATTCTTTGTTG	Variant confirmation	299	64	45
30	SYT16_JC007_R	CTTGCCTCTGGAAGAACCAG	Variant confirmation		64	55
31	KLC4_JC007_F	TGGGAAGTATGGGTTGAAGG	Variant confirmation	1050	60	50
32	KLC4_JC007_R	CTGCTGCAGACACTCAATGG	Variant confirmation		61	55
33	NUP188_JC007_F	TTTACTCATCTTTGCCTCTGG	Variant confirmation	525	57	43

S. No.	Primer name	Sequence (5'-3')	Application	Fragment size	Tm	%GC
34	NUP188_JC007_R	ACAGCTTCTGGGATTGTCAG	Variant confirmation		58	50
35	KIF23_JC002_F	TTTTTGCGTAACACATTTGG	Variant confirmation	273	57	35
36	KIF23_JC002_R	TTTCCCGACAAAAACTCAAG	Variant confirmation		58	40
37	KIF1A_FL_F_1	AAATGGGCGGTAGGCGTG	KIF1A construct and SDM confirmation	604	60	61
38	KIF1A_FL_R_1	CAGGTTGCCTTTGTTCTTGGG			60	52
39	KIF1A_FL_F_2	CCTCCTGAACCCCAAGAACAA	KIF1A construct and SDM confirmation	603	60	52
40	KIF1A_FL_R_2	TTGCAGCGGATCTGTTTGGC			62	55
41	KIF1A_FL_F_3	ATGATGAGACCCTCAGCACAC	KIF1A construct and SDM confirmation	603	60	52
42	KIF1A_FL_R_3	CAATGTCTTGCCGCCTCTCT			60	55
43	KIF1A_FL_F_4	GAGTGGGCAGAGAAGACGCA	KIF1A construct and SDM confirmation	682	62	60
44	KIF1A_FL_R_4	TTTAGCTCCACACTGATGGCA			60	48
45	KIF1A_FL_F_5	ACCAGTTCACTTCTCTTCGGG	KIF1A construct confirmation	603	60	52
46	KIF1A_FL_R_5	AGTGCTGGTCATCGAAGGAG			59	55
47	KIF1A_FL_F_6	ATGGATCAGGTGTCCGTCAG	KIF1A construct confirmation	672	59	55
48	KIF1A_FL_R_6	GCCGTGTCATTGTACTGAGC			59	55
49	KIF1A_FL_F_7	GGGTCATGCCGCTATCCAAG	KIF1A construct confirmation	611	61	60
50	KIF1A_FL_R_7	AAACAGGTTGCGAATGGAGC			59	50
51	KIF1A_FL_F_8	TTTTCTACTCCCGTGACGCC	KIF1A construct confirmation	799	60	55
52	KIF1A_FL_R_8	TCTGTCTCTGTTGCCCGAAC			60	55
53	KIF1A_FL_F_9	CCGAGCTCCTTCCAGAACTTG	KIF1A construct confirmation	605	61	57
54	KIF1A_FL_R_9	GCTGAACTTGTGGCCGTTT			59	53
55	KIF1A_FL_F_10	TCCAAGCTCTCCCGAAGGA	KIF1A construct confirmation	609	60	58
56	KIF1A_FL_R_10	CTCGATGTTGTGGCGGATCT			60	55
57	KIF1A_FL_F_11	TATCATGGCCGACAAGCAGA	KIF1A construct confirmation	729	59	50
58	KIF1A_FL_R_11	TCGCCCTGATAGACGGTTTTT			60	48
59	mCherry_F	CCCCGTAATGCAGAAGAAGA	KIF23 construct confirmation	NA	64	50
60	mCherry_R	TTGGTCACCTTCAGCTTGG			64	53
61	SV40_F	TATTTATGCAGAGGCCGAGG	KIF23 construct confirmation	NA	64	50
62	SV40_R	GAAATTTGTGATGCTATTGC			57	35
63	KIF1A_p.(Asp248Glu)_F	CAGCCTGGTGGAACTGGCGGGCAGT	SDM_Case 1 variant	NA	74	68
64	KIF1A_p.(Asp248Glu)_R	ACTGCCCGCCAGTTCCACCAGGCTG			74	68
65	KIF1A_p.(Cys92Arg)_F	CCATAGGCAAAGATGCGCACGTTGTAGCCCTCA	SDM_Case 2 variant	NA	79	55
66	KIF1A_p.(Cys92Arg)_R	TGAGGGCTACAACGTGCGCATCTTTGCCTATGG			79	55
67	KIF1A_p.(Pro305Leu)_F	ATACCGAGTCTCGGTAAAGGATGAAGTCTGTCTTC	SDM_Cases 3 and 12 variants	NA	78	46
68	KIF1A_p.(Pro305Leu)_R	GAAGACAGACTTCATCCTTTACCGAGACTCGGTAT			78	46

S. No.	Primer name	Sequence (5'-3')	Application	Fragment size	Tm	%GC
69	KIF1A_p.(Thr99Met)_F	TGCCAGCCCCTATCTGTCCATAGGCAAAGATGC	SDM_Cases 4 and 7 variants	NA	80	55
70	KIF1A_p.(Thr99Met)_R	GCATCTTTGCCTATGGACAGATAGGGGCTGGCA			80	55
71	KIF1A_p.(Tyr89Asp)_F	AAGATGCACACGTTGTCGCCCTCAAAGGCATGC	SDM_Case 5 variant	NA	76	55
72	KIF1A_p.(Tyr89Asp)_R	GCATGCCTTTGAGGGCGACAACGTGTGCATCTT			76	55
73	KIF1A_p.(Cys92Tyr)_F	GTCCATAGGCAAAGATGTACACGTTGTAGCCCTCA	SDM_Case 6 variant	NA	79	49
74	KIF1A_p.(Cys92Tyr)_R	TGAGGGCTACAACGTGTACATCTTTGCCTATGGAC			79	49
75	KIF1A_p.(Gly102Ser)_F	GTGTAGGACTTGCTAGCCCCTGTCTGTCCA	SDM_Case 8 variant	NA	79	57
76	KIF1A_p.(Gly102Ser)_R	TGGACAGACAGGGGCTAGCAAGTCCTACAC			79	57
77	KIF1A_p.(Gly117Val)_F	GAGCTGTGGGATGATGACTTGCTGGTCCTTCTC	SDM_Case 9 variant	NA	76	55
78	KIF1A_p.(Gly117Val)_R	GAGAAGGACCAGCAAGTCATCATCCCACAGCTC			76	55
79	KIF1A_p.(Ser217Tyr)_F	GAAGACGGCGTGATAACGGCTGCTGGTCT	SDM_Case 10 variant	NA	75	59
80	KIF1A_p.(Ser217Tyr)_R	AGACCAGCAGCCGTTATCACGCCGTCTTC			75	59
81	KIF1A_p.(Asp248Gly)_F	TGCCCGCCAGGCCCACCAGGCTG	SDM_Case 11 variant	NA	75	78
82	KIF1A_p.(Asp248Gly)_R	CAGCCTGGTGGGCCTGGCGGGCA			75	78
83	KIF1A_p.(Gly251Arg)_F	CTCGTTCACTGCGCGCCAGGTCCAC	SDM_Case 12 variant	NA	74	68
84	KIF1A_p.(Gly251Arg)_R	GTGGACCTGGCGCGCAGTGAACGAG			74	68
85	KIF1A_p.(Arg254Trp)_F	GTGGAGTCGGCCCATTCACTGCCCGCCAG	SDM_Cases 13 to 18 variants	NA	86	69
86	KIF1A_p.(Arg254Trp)_R	CTGGCGGGCAGTGAATGGGCCGACTCCAC			86	69
87	KIF1A_p.(Arg254Gln)_F	CGTGGAGTCGGCTTGTTCACTGCCCGC	SDM_Cases 19 and 20 variants	NA	83	67
88	KIF1A_p.(Arg254Gln)_R	GCGGGCAGTGAACAAGCCGACTCCACG			83	67
89	KIF1A_p.(Arg254Pro)_F	CGGGCAGTGAACCAGCCGACTCCAC	SDM_Case 21 variant	NA	80	68
90	KIF1A_p.(Arg254Pro)_R	GTGGAGTCGGCTGGTTCACTGCCCG			80	68
91	KIF1A_p.(Leu278Pro)_F	CAGAGATGACCTTTCCCGGGGTGGTCAAAGACTTG	SDM_Case 22 variant	NA	78	54
92	KIF1A_p.(Leu278Pro)_R	CAAGTCTTTGACCACCCCGGGAAAGGTCATCTCTG			78	54
93	KIF1A_p.(Arg307Pro)_F	ATGTCAATACCGAGTCTGGGTAAGGGATGAAGTCT	SDM_Cases 24 and 25 variants	NA	74	46
94	KIF1A_p.(Arg307Pro)_R	AGACTTCATCCCTTACCCAGACTCGGTATTGACAT			74	46
95	KIF1A_p.(Arg316Trp)_F	TTGACATGGCTTCTCTGGGAAAACCTGGGCG	SDM_Cases 26 to 30 variants	NA	79	55
96	KIF1A_p.(Arg316Trp)_R	CGCCCAGGTTTTCCCAGAGAAGCCATGTCAA			79	55
97	KIF1A_p.(Thr344Met)_F	GTCTGCATATCTCAGTATGCTGAGGGTCTCATCAT	SDM_Case 31 variant	NA	78	46
98	KIF1A_p.(Thr344Met)_R	ATGATGAGACCCTCAGCATACTGAGATATGCAGAC			78	46
99	KIF23_Case 1_F	CATCAAAGTAGTTCTTGAACACATGGGTTAACTTTGAATCTCGATA	SDM_KIF23_Case 1 variant	NA	77	35
100	KIF23_Case 1_R	TATCGAGATTCAAAGTTAACCCATGTGTTCAAGAACTACTTTGATG			77	35

Table A2: Online tools used in this thesis.

S. No.	Name of the tool	Web link	Application
1	ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/	Human variations and phenotypes
2	Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/	Multiple sequence alignment of sequences from different species
3.	CodonCode Aligner software	https://www.codoncode.com/aligner/	ABI sequence alignment and analysis
4	Combined Annotation Dependent Depletion (CADD)	https://cadd.gs.washington.edu/	*in silico* variant pathogenicity prediction
5	Cpipe	http://cpipeline.org/	NGS: variant detection
6	Decipher	https://decipher.sanger.ac.uk/	Clinical database to compare phenotypic and genotypic data
7	ExPASy - Translate tool	https://web.expasy.org/translate/	Translation of a nucleotide sequence to protein
8	Genome Aggregation Database (gnomAD)	http://gnomad.broadinstitute.org/	Mean allelic population frequency (genome sequencing)
9	Genomizer	http://genomizer.com/	Variant curation
10	HOPE (Centre for Molecular and Biomolecular Informatics)	http://www.cmbi.ru.nl/hope/	*in silico* protein modelling variant prediction tool
11	LOVDplus VCGS Research database	http://lovdresearch.mcri.edu.au/login	NGS: variant filtering
12	Matchmaker Exchange	https://www.matchmakerexchange.org/	Identification of individuals with genetic variations
13	Mutalyzer	https://www.mutalyzer.nl/	Sequence variant nomenclature
14	MutationTaster	http://www.mutationtaster.org/	*in silico* variant pathogenicity prediction
15	National Center for Biotechnology Information (NCBI)	https://www.ncbi.nlm.nih.gov/	Research publication and other literature
16	Online Mendelian Inheritance in Man (OMIM)	https://www.omim.org/	Gene information, clinical phenotype
17	Polyphen-2	http://genetics.bwh.harvard.edu/pph2/	*in silico* variant pathogenicity prediction
18	Primer3	http://bioinfo.ut.ee/primer3-0.4.0/	Primer design
19	PyMOL	http://www.pymol.org/	3D protein visualization software
20	QuikChange Primer Design	www.agilent.com/genomics/qcpd	Site directed mutagenesis primer design
21	RCSB Protein Data Bank (PDB)	http://www.rcsb.org/	3D protein structure
22	SIFT	http://sift.jcvi.org/	*in silico* variant pathogenicity prediction
23	SnapGene Viewer	https://www.snapgene.com/snapgene-viewer/	Construct sequence analysis
24	UCSC Genome Browser	https://genome.ucsc.edu/	Gene sequence information

Appendix B

Construct maps

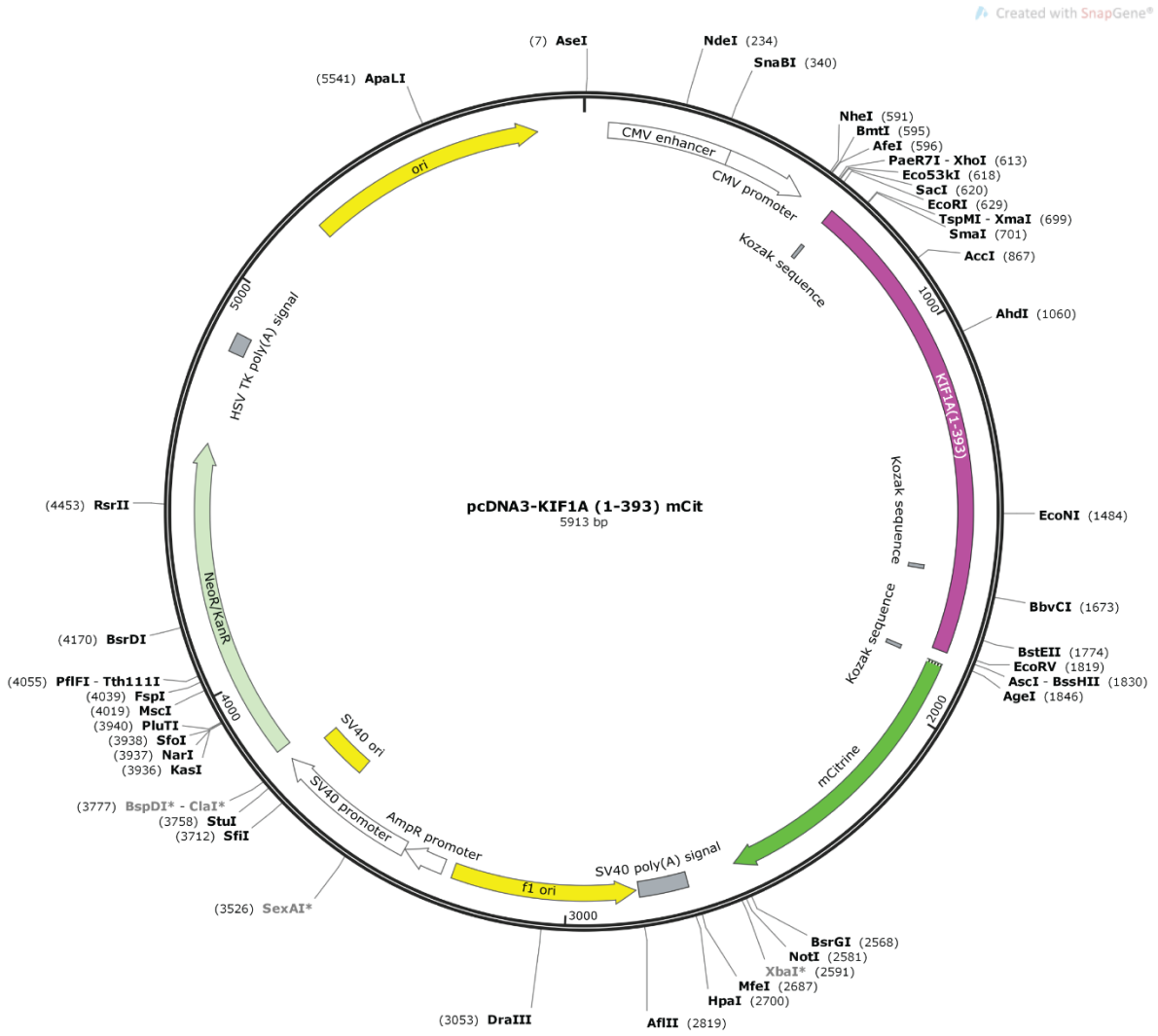


Figure B1: Map of the KIF1A(1-393)-3xmCit construct used for the neurite tip accumulation assay in the *KIF1A* studies.

This construct contains sequence that encodes for dimeric KIF1A motor (amino acid 1-393; pink) consisting of motor domain (amino acids 1-361), neck linker and neck coil domains which is dimerized via GCN4 leucine zipper (Hammond et al., 2009). Three tandem copies of mCitrine (mCit) labels dimeric KIF1A motor protein with a fluorescent signal. The total vector size: 5913 bp, backbone size without insert: 4734 bp, insert size: 1179 bp, vector type: mammalian expression.

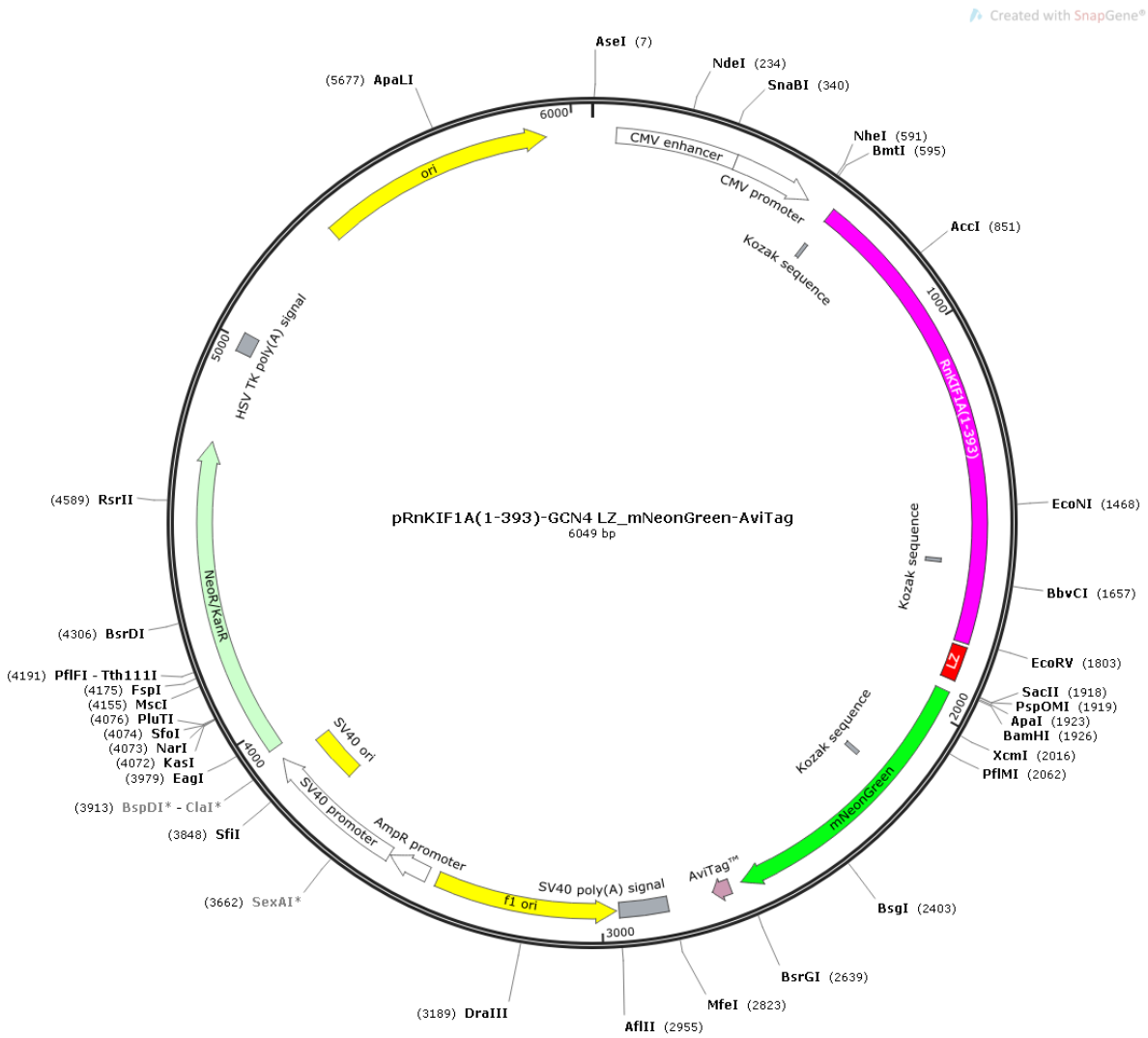


Figure B2: Map of the KIF1A(1-393)-LZ-mNG-AviTag construct used for the *in-vitro* microtubule gliding assay in the *KIF1A* studies.

This encodes for dimeric KIF1A motor (amino acid 1-393; pink) containing motor domain (amino acids 1-361), neck linker and neck coil domains which gets dimerized via GCN4 leucine zipper (Hammond et al., 2009). Fluorescent mNeonGreen (mNG) is fused at the C-terminal. A second tag at the C-terminal, AviTag (amino acid sequence GLNDIFEAQKIEWHE), attaches the biotin label to the expressed protein when co-expressed with HA-tag labelled bacterial biotin ligase BirA (Yue et al., 2018).

Figure B3: Map of the pmCherry-C1-MKLP1 construct used for the neurite tip accumulation studies in the *KIF23* studies.

This construct (7274 bp, plasmid number: 70154; Addgene) was used for the mammalian expression of MKLP1/ KIF23 (orange) with a mCherry (red) fusion (Joseph et al., 2012). The total vector size: 7274 bp, backbone size without insert: 4722 bp, insert size: 2574bp, vector type: mammalian expression.

Appendix C

Videos for *in-vitro* microtubule gliding assay

An example of raw TIRF videos, Video C1 to Video C5, for *in-vitro* microtubule gliding assay has been uploaded on CloudStor (https://cloudstor.aarnet.edu.au/plus/s/ANYMiyGFHf4qPEk)

Video C1: Example of a recording of the *in vitro* microtubule gliding assay for the wild type protein.

Video C2: Example of a recording of the *in vitro* microtubule gliding assay for the p.(Asp248Glu) protein.

Video C3: Example of a recording of the *in vitro* microtubule gliding assay for the p.(Cys92Arg) protein.

Video C4: Example of a recording of the *in vitro* microtubule gliding assay for the p.(Pro305Leu) protein.

Video C5: Example of a recording of the *in vitro* microtubule gliding assay for the p.(Arg316Trp) protein.

Appendix D

Reviewer's comments for publication in Chapter 3

The following reviewer's comments have been addressed and the manuscript is now in second round of review.

Referee: 1 - Comments to the Author

This is an interesting paper that sets out clinical and molecular information about four cases with (pathogenic) variants in KIF1A and functional studies that demonstrate a plausibly causal role of these variants in the neurological dysfunction.

My one concern is with the effort to relate the clinical phenotypes to Rett syndrome, as I do not find that fully convincing. I have additional questions that, I feel, need to be answered about all four cases. The approach taken in the paper appears to me (as a clinician) to be a bit too simplistic, as if the authors have adopted a rather uncritical tick-box approach to the diagnosis.

1. In Cases 1 and 4, it would be helpful to know more about the onset of seizures and their possible relationship to the loss of acquired skills, especially in Case 4. If regression does seem to have coincided with refractory seizures, then the association with Rett syndrome would essentially disappear.

   Response:

   Seizures occur in about 80% of RTT individuals and its onset is age dependent and more common in individuals with severe early developmental delay. Usually, the onset of seizures is around four years of age which is later as compare to other neurological disorders (Jian et al., 2007).

   Case 1: This girl had static development from around 12 months and then was noted to have global development delay from around 15 months of age. She had developmental regression from 16 months with loss of communication skills and then loss of hand skills. She had seizures from 3 years of age treated with carbamazepine. Hence, the

developmental regression pre-dated the onset of seizures. She had uncontrolled seizures between 3-5 years of age. We have updated the clinical story in the supplementary material accordingly.

Case 2: She developed atonic seizures at 4 years of age and was started on clonazepam with partial control of the seizures. We have updated the information in under detailed case report section in supplementary material.

Case 3: He has slow paroxistic bilateral posterior activity in sleep. Repeated EEGs showed abnormal activity without seizures.

Case 4: Her presenting diagnosis was optic nerve atrophy and probable cortical visual impairment with severe encephalopathy. The loss of skills at six months old occurred prior to her first diagnosed seizure at one year old, making it less likely this loss was primarily epileptic.

2. Assuming that the loss of skills was not simply seizure-related, I wonder why Case 4 was not regarded as having Classic Rett Syndrome, as it is not a requirement for the diagnosis of RTT to lose skills that have not been acquired. However, one would expect the loss of pre-verbal vocalisation or a withdrawal of social interest. So additional clinical information is required.

Response:

During initial interview with a member of our research staff, proband's mother indicated loss of acquired purposeful hand skills. When we asked for further clarification, it turns out the mother may have misunderstood the initial question. The proband never had purposeful hand skills. She had recurrent mouthing and chewing of left hand, sometimes biting so hard that she would injure herself. This started ~2-3years of age, and continued until her death. This repetitive activity cannot be regarded as stereotypic central hand movements that are often seen in RTT individuals. Thus, based upon recent information, case 4 can be re-classified as RTT-like as she exhibited a period of regression and fulfilled one main and

five supportive criteria. We have updated the clinical story in Table 1 in main text file and also in supplementary material accordingly.

3. Did Case 1 have any perinatal events that could explain the MRI scan anomalies? That is clearly crucial, as it would lead one to remove the label of Rett syndrome altogether.

   Response:
   No. The pregnancy, delivery and perinatal period were all normal.

4. In both cases 2 and 3, without any suggestion of regression, I wonder why the idea of Rett syndrome has appeared at all. The features in common with Rett syndrome are too few to really warrant any mention of RTT at all. If these case are included, then most cases of non-progressive cognitive impairment could also be regarded as Rett-like. Why use that term at all? The term is NOT used in the Neul et al paper on the revised diagnostic criteria.

   Response:
   Over the last few years, a new classification, known as RTT-like, is slowly evolving, that represents a large proportion of cases who exhibit a number of clinical features of RTT, but not enough in order to be classified as either classic or atypical RTT. Even though there is no formal definition of RTT-like cases, an increasing number of genes are being associated with clinical phenotypes overlapping RTT features, along with features characteristic of other neurodevelopmental disorders, and so the genetic landscape of RTT and RTT-like individuals is constantly evolving (Iwama et al., 2019).

   Case 2 did not have a period of normal development followed by regression, and in fact had grossly abnormal development in the first 6 months. So, this excludes her from falling under the classification of classic or atypical RTT. However, she otherwise has a number of the main (2/4) and supportive criteria (2/11) and thus we may classify her as RTT-like individual. Case 3 also did not have period of regression but he fulfils several main and supportive criteria, and so has also been designated as RTT-like.

5. Finally, we need some further information about case 2 and her diagnosis of MCAD deficiency. Did she have any episodes of metabolic decompensation as an infant? Have further biochemical or metabolic studies been carried out?

Response:

She did not undergo newborn screening for medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, but had a biochemical diagnosis of medium-chain acyl-CoA dehydrogenase deficiency based on persistently elevated plasma C6, C8, C10 and C10:1 acylcarnitine levels. She was heterozygous for the common known c.985A>G (p.(Lys304Glu) pathogenic mutation in *ACADM* gene. A variant in the second allele was not identified, and so far she has not had episodes of hypoglycemia or metabolic decompensation. So, whilst it is possible that she is merely a carrier of *ACADM* carrier, the persistent biochemical abnormalities mandate that she should be treated as if she has MCAD deficiency. Thus, she has been on carnitine therapy since the age of two years. We have updated the clinical story in the supplementary material accordingly.

In summary, there is no real connection to Rett syndrome in cases 2 or 3 and some grounds to doubt it in the other two cases as well. It would be much more appropriate - to my mind - to report these cases as extending the phenotype of KIF1A-related disorders and to mention in the discussion if any of them really has a phenotype that could be misinterpreted as corresponding to Rett syndrome (but that would require additional clinical information to be at all convincing).

Response:

We hope that the additional information that we have provided for cases 1 and 4 is more convincing. Based on our experience we are confident that the clinical diagnosis of Rett syndrome is solid in case 1. In addition, cases 2, 3 and 4 have features beyond what has been previously recognised for KIF1A-related disorders, and fit within the Rett-like spectrum. So, we are indeed in agreement that for these three cases, the Rett-like features extend the phenotype of KIF1A-related disorders. The key message we want to convey is that when encountering a child with Rett-like phenotype or outright RTT, but who is

negative for *MECP2* mutations, consideration should be given to screening the *KIF1A* gene.

Referee: 2 - Comments to the Author

The manuscript "Phenotypic spectrum of de novo missense variants in kinesin family member 1A (KIF1A): expansion to include Rett syndrome and Rett-like individuals" by Kaur et al. is a very well written and interesting work showing for the first time a link between mutations in the microtubule motor protein KIF1A and Rett syndrome. The authors identify 4 different missense mutations in KIF1A in individuals with Rett syndrome or Rett-syndrome like features and analyse the effect of the mutations in two functional assays.
I only have minor points that I suggest the authors to address to render the work ready for publication in Human Mutation.

1. Introduction: whereas both MECP2 and CDKL5 regulate microtubule functions no data describe a common role in regulating microtubule stability.

   Response: We have updated the main text of the file on page 4 and added supporting references.

   Before revision: There is clear evidence that the RTT-associated genes *MECP2* and *CDKL5* share a common role in regulating microtubule stability and neuronal transport (Table S1).

   After revision: There is clear evidence that the RTT-associated genes *MECP2* and *CDKL5* share a common role in regulating microtubule stability and neuronal transport, and that pathogenic mutations in those genes disrupt microtubule dynamics (Barbiero et al., 2019; Delepine et al., 2013; Gold et al., 2015).

2. A) Does the expression of KIF1A derivatives in SH-SY5Y cells cause any alteration in neurite outgrowth during RA-induced differentiation?

Response: For the neurite tip accumulation assay, several parameters were optimised, including the concentration of retinoic acid (RA) to be used for induction of differentiation in SH-SY5Y cells. Of the three different chosen concentrations of RA (5µM, 10µM and 20µM), optimum differentiation with low toxicity over a seven-day period was observed when 10µM of RA was used to differentiate un-transfected SH-SY5Y cells. We included untransfected control SH-SY5Y cells in all repeat neurite tip accumulation assay experiments, and we observed no qualitative differences in neurite outgrowth in transfected versus untransfected SH-SY5Y cells differentiated with 10µM of RA over a seven-day time course.

To address the reviewer's suggestion, we have updated the main text of the file.

Before revision: After 24 hours the media was refreshed to RPMI with 1% serum supplemented with 10µM retinoic acid (RA; Sigma; R2625) to induce neuronal differentiation and the media was refreshed daily for 4 days.
After revision: After 24 hours the media was refreshed to RPMI with 1% serum supplemented with 10µM retinoic acid (RA; Sigma; R2625) to induce neuronal differentiation and the media was refreshed daily for 4 days. The expression of KIF1A motor protein in SH-SY5Y cells did not appear to cause any qualitative alteration in neurite outgrowth during RA-induced differentiation (data not shown).

B) The authors should describe more precisely how they evaluate the accumulation of the expressed protein at the neurite tip: do they just evaluate the capacity of the protein to reach the neurite tip or to be preferentially located at the tip. The WT protein indeed seems to preferentially accumulate at the tip whereas the Case 1 is absent from the tip and the Case 2 derivative seems to be equally distributed.

Response: As might be expected, the different variants studied showed variability in their localization pattern along the length of neurite after differentiation. So, during analysis using ImageJ, the mean fluorescence intensity (MFI) along the whole length of the growing neurite was calculated manually from the projection point of cell body where the neurite begins. MFI at the cell body was also calculated manually by drawing around the boundary

of the cell body. Hence, using the neurite tip accumulation assay in this study, the capacity of the expressed protein to reach the neurite tip is being assessed.

To address the reviewer's suggestion, we have updated the main text of the file on page 8 and other relevant sections. Please see the track changes in the main text file to view all the appropriate changes made.

Before revision: The quantification of the mean florescence intensity (MFI) at the neurite tip and cell body of SH-SY5Y cells was measured manually using ImageJ.

After revision:
Quantification of the mean florescence intensity (MFI) along the length of the neurites and within the cell bodies of SH-SY5Y cells was measured manually using ImageJ. In addition, the y-axis of the graph in figure 2 has been modified to "MFI neurite length: cell body".

3. Microtubule gliding assay: it should be described how the authors observe the lack of microtubule binding. It is not clear.

A unique property of the total internal reflection fluorescence (TIRF) microscopy is that it captures the signal from fluorescent microtubules in a limited specimen region immediately adjacent to the adsorbed KIF1A motor protein and the glass coverslip. This exclusive property minimizes any noise coming from the floating microtubules that were not able to bind to the KIF1A motor protein. Thus, the lack of the microtubule binding was mainly assessed qualitatively by the observation that a significantly smaller proportion of fluorescent labelled microtubules were found to bind to the adsorbed variant, in particular the [p.(Cys92Arg) and p.(Pro305Leu)] KIF1A motor proteins, compared to the wild type KIF1A motor protein at the same interface optimised using wild type protein.

Moreover, in the raw videos, the ends of the partially bound microtubules for p.(Cys92Arg) and p.(Pro305Leu) variants were observed to be flickering before dissociating from the adsorbed protein after a short period of time and moving out of the pre-set interface for

TIRF. This phenomenon during the gliding assay suggests that there was impaired microtubule binding.

Before revision:

3.2.2 In-vitro microtubule gliding assay reveals impaired microtubule binding

To directly examine the ability of individual variant KIF1A motor proteins to undergo microtubule-based motility, we performed in-vitro TIRF-based microtubule gliding assays. In this assay, truncated KIF1A(1-393) motors are immobilized on a slide and their ability to work in concert to "glide" microtubules along the surface is examined by fluorescence microscopy. For the wild-type KIF1A(1-393) motors, the majority of microtubules (~85-90%) introduced into the flow chamber bound to the immobilized KIF1A and were moved along the surface with an average velocity of 0.91 ± 0.09 μm (Figure 3 and Video 1).

In the case of the individual variant [p.(Asp248Glu)] in switch II which is predicted to impact the coordination of the divalent cation and nucleotide, the majority of microtubules introduced into the chamber (~70%) were able to bind to the variant KIF1A motors but moved at a significantly lower velocity (0.005 ± 0.002 μm/s) compared to the wild-type motor protein (Figure 3 and Video 2). For the p.(Cys92Arg) and p.(Pro305Leu) variants, very few of the microtubules introduced to the chamber were able to bind to the variant KIF1A motors (~10%) and the microtubules that did bind were observed to dissociate after a short period of time, consistent with the hypothesis that these variants alter protein folding and/or microtubule binding.

After revision:

3.2.2 In-vitro microtubule gliding assay reveals impaired microtubule binding

To directly examine the ability of individual variant KIF1A motor proteins to undergo microtubule-based motility, we performed in-vitro TIRF-based microtubule gliding assays. In this assay, truncated KIF1A(1-393) motors are immobilized on a slide and their ability to work in concert to "glide" microtubules along the surface is examined by fluorescence microscopy. A unique property of the TIRF microscopy is that it captures the signal from fluorescent microtubules in a limited specimen region immediately adjacent to the adsorbed KIF1A motor protein and the glass coverslip. This exclusive property minimizes any noise

coming from the floating microtubules that were not able to bind to the KIF1A motor protein.

For the wild-type KIF1A(1-393) motors, the majority of microtubules (~85-90%) introduced into the flow chamber bound to the immobilized KIF1A and moved along the surface with an average velocity of 0.91 ± 0.09 μm (Figure 3, Supp. Figure S4and Supp. Video 1).

For the individual variant [p.(Asp248Glu)] in switch II, which is predicted to impact the coordination of the divalent cation and ATP, the majority of microtubules introduced into the chamber (~70%) were able to bind to the variant KIF1A motors but moved at a significantly lower velocity (0.005 ± 0.002 μm/s) compared to the wild-type motor protein (Figure 3, Supp. Figure S4 and Supp. Video 2). For the p.(Cys92Arg) and p.(Pro305Leu) variants, very few of the microtubules introduced to the chamber were able to bind to the variant KIF1A motors (~10%) (Figure 3, Supp. Figure S4 and Supp. Video 3, 4). The lack of the microtubule binding was mainly assessed qualitatively by the observation that a significantly smaller proportion of fluorescently labelled microtubules were found to bind to the adsorbed variant as compared to the wild type KIF1A motor protein at the same TIRF interface optimized using wild type protein. Moreover, in the raw videos, the ends of the partially bound microtubules for the p.(Cys92Arg) and p.(Pro305Leu) variants were observed to be flickering before dissociating from the adsorbed protein after a short period of time and moving out of the pre-set interface for TIRF. This phenomenon during the gliding assay suggests that there was impaired microtubule binding, consistent with the hypothesis that these variants alter protein folding and/or microtubule binding.

Appendix E

Detailed clinical information of KAND individuals

Table E1: Detailed clinical phenotype of 27 individuals with KAND

Clinical feature	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Gender	M		M	F	M	M	F	F	M	F	F	M	F	nr	F	F	M	M	F	F	F	M	F	F	F	F	F
Neurological issues																											
Hypotonia	-	-	+	-	+	+	-	+	+	+	+	+	+	+	-	-	+	+	+	-	+	+	-	+	+	-	-
Hypertonia	+	+	-	+	-	-	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+	+	-	+	+	-
Microcephaly	-	+	+	+	+	-	-	+	-	+	-	-	+	-	-	-	-	-	-	-	-	-	-	-	+	+	-
Global DD	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
ID	+	nr	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	nr	+
Ataxia	nr	-	nr	-	nr	nr	nr	nr	nr	nr	+	nr	+	+	nr	+	nr	nr	nr	nr	nr	+	+	+	+	nr	+
Spastic paraplegia	nr	+	nr	+	nr	nr	nr	nr	nr	nr	nr	nr	+	+	nr	nr	nr	nr	nr	nr	nr	+	+	+	+	+	+
Epilepsy	-	+	-	-	+	-	+	-	+	-	+	+	+	-	-	-	-	-	+	+	+	+	+	+	-	-	-
Peripheral neuropathy	nr	-	nr	-	nr	+	nr	nr	nr	nr	+	nr	-	-	+	+	nr	nr	nr	+	+	+	-	-	-	nr	-
Nerve conduction studies	nr	nr	nr	-	nr	nr	nr	nr	nr	nr	nr	nr	-	-	nr	nr	nr	nr	nr	nr	nr	+	nr	nr	nr	nr	-
Seizures	-	-	+	-	-	+	-	+	-	+	-	+	+	+	-	-	-	-	-	+	+	-	+	+	-	nr	-
MRI changes																											
Cerebral atrophy	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	nr	-	-	-	-	nr	-	-	nr	nr
Cerebellar atrophy	-	-	nr	-	-	+	-	-	+	+	+	+	+	+	-	-	-	nr	-	+	+	+	nr	+	-	nr	nr
Reduced cerebral white matter	-	-	-	-	-	-	-	-	-	nr	nr	nr	+	-	-	-	-	nr	-	nr	nr	-	nr	-	-	nr	nr
Thin/ absent corpus callosum	-	-	+	-	-	-	-	+	-	-	+	-	-	-	-	-	-	nr	-	-	-	-	nr	-	-	nr	nr
Muscle issues																											

Clinical feature	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Upper limb spasticity	nr	+	nr	-	nr	nr	nr	nr	nr	nr	nr	nr	nr	+	nr	nr	nr	nr	nr	nr	nr	+	-	-	-	nr	-
Lower limb spasticity	nr	+	nr	+	nr	nr	nr	nr	nr	nr	nr	nr	nr	+	nr	nr	nr	nr	nr	nr	nr	+	+	+	+	nr	+
Muscle hypertrophy	-	-	-	-	-	-	-	-	-	nr	nr	nr	nr	-	-	nr	-	-	-	-	-	-	-	-	-	nr	-
Muscle weakness	nr	+	nr	+	nr	nr	nr	nr	nr	nr	nr	nr	nr	+	nr	nr	nr	nr	nr	nr	nr	-	-	-	-	nr	+
Gait abnormalities	nr	+	+	+	nr	+	+	+	+	nr	nr	nr	nr	+	+	+	+	+	+	+	+	+	+	+	+	nr	+
Ophthalmological Issues																											
Cataracts	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	nr	-
Cortical visual impairment	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	nr	-
Optic nerve abnormalities	-	-	+	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	+	+	+	nr	nr	+	nr	+
Ptosis	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	nr	-
Strabismus	-	-	+	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	+	+	+	nr	-
Others																											
Gastroesophageal reflux disease	-	-	+	-	-	+	+	+	-	+	-	-	-	-	+	-	-	-	-	+	+	-	-	-	-		-
Eczema	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-	-	+	+	-	-	+	+	-	-	-		-
Genital abnormalities	-	-	+	-	-	+	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-		-

nr = not reported; + = present; - = absent

Appendix F

Submission proof for publication in Chapter 5

From: eesserver@eesmail.elsevier.com on behalf of Journal of Genetics and Genomics
To: Simran Kaur
Subject: Track your recent Co-Authored submission to JGG
Date: Thursday, 13 February 2020 7:51:17 AM

*** Automated email sent by the system ***

Dear Dr. Simranpreet Kaur,

You have been listed as a Co-Author of the following submission:

Journal: Journal of Genetics and Genomics
Title: Expanding the genetic landscape of Rett syndrome to include Lysine (K) acetyltransferase 6A (KAT6A)
Corresponding Author: John Christodoulou
Co-Authors: Simranpreet Kaur; Nicole J Van Bergen; Bruria Ben-Zeev; Emanuela Leonardi; Tiong Y Tan; David Coman; Benjamin Kamien; Susan M White; Miya St John; Dean Phelan; Kristin Rigbye; Sze Chern Lim; Michelle C Torres; Melanie Marty; Elena Savva; Teresa Zhao; Sean Massey; Alessandra Murgia; Wendy A Gold;

To be kept informed of the status of your submission, register or log in (if you already have an Elsevier profile).

Register here: https://ees.elsevier.com/jgg/default.asp?acw=&pg=preRegistration.asp&user=coauthor&fname=Simranpreet&lname=Kaur&email=simran.kaur@mcri.edu.au

Or log in: https://ees.elsevier.com/jgg/default.asp?acw=&pg=login.asp&email=simran.kaur@mcri.edu.au

If you did not co-author this submission, please do not follow the above link but instead contact the Corresponding Author of this submission at john.christodoulou@mcri.edu.au.

Thank you,

Journal of Genetics and Genomics

Appendix G

Additional publications referenced in Chapter 6

Publication 1

Schonewolf-Greulich, B., A. M. Bisgaard, M. Duno, C. Jespersgaard, M. Rokkjaer, L. K. Hansen, E. Tsoutsou, C. Sofokleous, M. Topcu, **S. Kaur**, N. J. Van Bergen, K. Brondum-Nielsen, M. J. Larsen, K. P. Sorensen, J. Christodoulou, C. R. Fagerberg and Z. Tumer (2019). Mosaic *MECP2* variants in males with classical Rett syndrome features, including stereotypical hand movements. *Clin Genet, 95*(3), 403-408. doi:10.1111/cge.13473

Received: 29 August 2018 | Revised: 5 November 2018 | Accepted: 6 November 2018

DOI: 10.1111/cge.13473

WILEY CLINICAL GENETICS

SHORT REPORT

Mosaic *MECP2* variants in males with classical Rett syndrome features, including stereotypical hand movements

Bitten Schönewolf-Greulich[1,2] | Anne-Marie Bisgaard[1] | Morten Dunø[3] | Cathrine Jespersgaard[2] | Mette Rokkjær[4] | Lars K. Hansen[5] | Eirini Tsoutsou[6] | Christalena Sofokleous[6] | Meral Topcu[7] | Simran Kaur[8] | Nicole J. Van Bergen[8] | Karen Brøndum-Nielsen[3] | Martin J. Larsen[9] | Kristina P. Sørensen[9] | John Christodoulou[8] | Christina R. Fagerberg[9] | Zeynep Tümer[2,10]

[1]Department of Paediatrics and Adolescent Medicine, Center for Rett Syndrome, Rigshospitalet, Copenhagen, Denmark

[2]Kennedy Center, Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[3]Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[4]Department of Pediatrics, Kolding Hospital, Kolding, Denmark

[5]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark

[6]Medical Genetics Department, National and Kapodistrian University of Athens, "Aghia Sophia" Children's Hospital, Athens, Greece

[7]Department of Medical Genetics, Hacettepe Medical School, Ankara, Turkey

[8]Department of Paediatrics, Murdoch Children's Research Institute, University of Melbourne, Melbourne, Australia

[9]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark

[10]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Correspondence
Bitten Schönewolf-Greulich, Center for Rett Syndrome, Kennedy Center, Rigshospitalet, Gl. Landevej 7, 2600 Copenhagen, Denmark.
Email: bitten.schoenewolf-greulich.01@regionh.dk

Funding information
Dronning Louises Børnehospitals Forskningsfond; the Victorian Government's Operational Infrastructure Support Program; Victorian Government

Rett syndrome is rarely suspected in males because of the X-linked dominant inheritance. In the literature, only six male patients have been reported with methyl-CpG-binding protein 2 (*MECP2*) mosaicism. Next-generation sequencing (NGS) methods have enabled better detection of somatic mosaicism compared to conventional Sanger sequencing; however, mosaics can still be difficult to detect. We present clinical and molecular findings in two males mosaic for a pathogenic *MECP2* variant. Both have been reexamined using deep sequencing of DNA isolated from four different cell tissues (blood, muscle, fibroblasts and oral mucosa). Deep sequencing of the different tissues revealed that the variants were present in all tissues. In one patient, the molecular diagnosis could only be established by reexamination after a normal whole exome sequencing, and the other case is an example of reverse genetic diagnostics. Rett syndrome should be considered in males with neurodevelopmental delay and stereotypical hand movements. Subsequent to clinical diagnosis males should be investigated with NGS-based technologies of *MECP2* with high read depth and a low threshold for variant calls. If the initial analysis on full blood derived DNA fails to confirm the suspicion, we recommend repeating the analysis on another tissue, preferentially fibroblasts to increase the diagnostic yield.

KEYWORDS

male, MECP2, mosaicism, NGS, Rett syndrome

1 | INTRODUCTION

Rett syndrome (RTT) (OMIM# 312750) is a neurodevelopmental syndrome first described in 1966 by Andreas Rett, who investigated a group of girls with regression of cognitive and motor function especially of speech and hand use and with development of hand stereotypies.[1] The latest international diagnostic criteria published by Neul et al[2] are being used to establish the clinical diagnosis of RTT. The primary causative gene is the X-linked gene methyl-CpG-binding protein 2 (*MECP2*).[3] Around 3900 females with *MECP2* variants have been reported in the international RettBASE.[4] In males, however, variants are not so common and only about 60 cases have been reported. The phenotype of males with single nucleotide variants is often more severe with neonatal encephalopathy as the main symptom in contrast to the classical RTT phenotype described in the majority of females.[5] The RTT phenotype observed in the few published male cases is occasionally explained by the presence of mosaicism[6–11] or a coincidental 47,XXY karyotype detected with Sanger sequencing of *MECP2* or chromosome analysis, respectively, using DNA derived from blood.[12,13] The six published cases of males with *MECP2* mosaicism have been diagnosed with non-syndromic neurodevelopmental delay, Rett-like, atypical RTT or classical RTT.[6–11] None of these patients were evaluated according to the 2010 criteria.[2] Recently, the unaffected father of a RTT patient hemizygous for c.925C>T (p.Arg309Trp) in *MECP2* was found to be mosaic for this variant.[14] This variant is, however, also described as resulting in a less severe phenotype than RTT and was also detected in an unaffected mother in non-mosaic state.[15]

Today, next-generation sequencing (NGS) methods, such as genome sequencing, exome sequencing or targeted gene panels, are commonly used as a first-tier approach to screen for genomic variants causing neurodevelopmental syndromes, including RTT. In the present study, two males with neurodevelopmental syndromes and mosaic for an *MECP2* variant are described. Only one of them was suspected of having RTT initially without finding a variant in the known RTT genes. The six previously published cases were also reevaluated according to the 2010 guidelines[2] either by using the information published or by personal communication with the authors, and the symptoms were compared with the two present cases.

2 | MATERIALS AND METHODS

2.1 | Subjects

In this study, two male patients are included. Patient A is an 8-year-old male (Figure 1A1, A2), patient B is a 10-year-old male (Figure 1B1, B2). The phenotypic features of both patients are listed in Table 1 (elaborated subject description in Supporting Information File A).

2.2 | Genetic study

The two clinical cases were examined as part of a project investigating patients with Rett-like syndromes: "Rett-like childhood syndromes – identification of disease causing genes and mechanisms, a clinical and molecular study." The project was approved by the Danish regional ethics committee and the regional Danish data protection agency, and the families have consented to the publication of photos and clinical information.

Apart from the thorough clinical description of the present cases, we also contacted the authors of the previously described males with mosaic *MECP2* variants in order to update and expand their previously published clinical presentations. Further information was gathered for two of the cases,[8,9] and features of the remaining cases are extracted from the publications.[6,7,10,11]

DNA was extracted from blood, buccal swab, skin and muscle biopsies using standard methods. To investigate the overall presence of the variant allele fraction (VAF) in different tissues, all DNA samples from the two patients were sequenced with the same method with high coverage.

Polymerase chain reaction amplification of *MECP2* exon 4 was performed, and the resulting products were fragmented. NGS was carried out on the IonProton (ThermoFisher, Waltham, Massachusetts) using standard procedures and settings (TorrentSuite v5.4, Waltham, Massachusetts, USA). All samples were sequenced in a single run. The sequencing depth varied from 19.428 to 39.344 reads.

Post hoc Sanger sequencing was carried out for the same samples in order to see whether the variant was detectable by this method. Parental/guardian consent was obtained.

3 | RESULTS

Table 1 summarizes the clinical features of the present cases together with the features of the previously published male patients with mosaic *MECP2* variants.[6–11] The published cases were diagnosed with non-syndromic neurodevelopmental delay, Rett-like, atypical RTT or classical RTT. Re-evaluation of the published cases according to the Neul 2010 criteria[2] suggest that all the patients have classical RTT, regardless of what the initial diagnosis was.

DNA extracted from blood, buccal swab, fibroblasts and muscle biopsies from both patients were sequenced. The resulting VAF values are shown in Table 1. Patient A displayed a VAF of 8.4% to 15.8%, with the highest VAF found in muscle and the lowest in blood (Table 1). In patient B, the VAF varied from 33% to 45%, with the highest VAF found in fibroblasts and the lowest in oral mucosa.

Post hoc Sanger sequencing showed that the variant in patient A was difficult to assess compared to patient B (Supplementary file B).

4 | DISCUSSION

We present two males with a mosaic *MECP2* variant. The RTT diagnosis was not even considered for patient B prior to exome sequencing at the age of 9 years, while for patient A the suspected clinical RTT diagnosis could not be confirmed by Sanger or initial exome sequencing. Among the more than 100 patients with a constitutional *MECP2* variant registered at the Danish National Center for Rett syndrome[16] between 2001 and 2016, there is only one male patient.[17] Our finding of two male patients in a relatively short period suggests that RTT in

FIGURE 1 A1, Patient A 1 years old using his hands. A2, Patient A 8 years old with hand stereotypies. B1, Patient B 4 years old still able to sit independently. B2, Patient B 9 years old he sits with support and has almost constant hand stereotypies [Colour figure can be viewed at wileyonlinelibrary.com]

males may be underdiagnosed. Males with a pathogenic constitutional *MECP2* variant are rare probably because of a severe phenotypic outcome, but in a mosaic state the phenotypic effect has been suggested to resemble that of females with RTT.[5] When reviewing the clinical data of the eight males with a mosaic *MECP2* variant, including the present cases it is clear that they all present with a classical RTT phenotype: a regression period with loss of skills and thus poor hand function, an abnormal gait and poor speech development. Finally, they all have a very recognizable feature of RTT, namely stereotypic hand movements (Figure 1). However, their diagnosis of RTT was delayed in most of the cases (2-6 years) presumably because of the male gender[11] or not even considered (Patient B) (Table 1); whereas the corresponding mean age of the diagnosis in females with RTT in Denmark is 2.5 years.[18]

Aside from the importance of raising clinicians' awareness of males with RTT, the present cases also illustrate the pitfalls in the diagnosis of mosaicism. Sanger sequencing is not the method of choice for detection of mosaicism. NGS-based methods necessitate high sequencing depth (optimal > 100× coverage) to reliably detect low-frequent mosaic variants. Diagnostic exome sequencing analyses are often carried out at lower read depth (approximately 30×),[19] and some regions are captured poorly; therefore, a targeted gene panel with high coverage of the gene of interest (eg, *MECP2*) and high read depth may be a preferred approach for the molecular genetic diagnosis of mosaic males who have typical or atypical RTT. The decision to select the most relevant approach can be reached through collaboration between clinicians and molecular geneticists to enhance the diagnostic rate of RTT in male patients.

The mosaic variants probably arise very early in embryogenesis and therefore most likely are present in all types of tissues. However, there can be asymmetry in the selective pressure among cell-lines,[20] resulting in different VAFs in different tissues. Standard genetic analyses are performed in DNA from blood samples and although mosaicism will probably be present in all cell types and not restricted to a single tissue, it is recommended to also test another tissue whenever mosaicism is suspected.[20]

TABLE 1 Clinical and molecular data of male cases with mosaicism

	Present cases		Published cases with new information		Previously published cases without new information			
	Case A	Case B	Psoni et al[9]	Topcu et al[8]	Pieras et al[10]	Kleefstra et al[11]	Clayton-Smith et al[6]	Armstrong et al[7a]
Gender	Male	Male	Male	Male	Male	Male	Male	Male
Age at evaluation (years)	8	9	14	12	4	11	6	14
RTT diagnosis considered at age	4 years	Never considered due to male gender and facial expression	6 years	2 years	3 years	Considered but disregarded due to male gender	6 years	uk
Head circumference	50 cm (<10th centile)	50.5 cm (<10th centile)	54.5 cm (50th centile)	47 cm (25th centile)	50 cm (50th centile)	51.5 cm (10th centile)	<25th centile	uk
Nucleotide change (NM_004992.3)	c.1308dupT	c.808C>T	c.316C>T	c.808C>T	c.360T>G	c.473C>T	c.241_242del	c.398G>A
Predicted effect on protein sequence	p.(Gln437Serfs*50)	p.(Arg270*)	p.(Arg106Trp)	p.(Arg270*)	p.(Tyr120*)	p.(Thr158Met)	p.(Gly81Glnfs*9)	p.(Arg133His)
Detected in								
Leukocytes	9%	36%	Mosaic	36%	Mosaic	Mosaic	Mosaic	Mosaic
Muscle tissue	16%	40%	nt	nt	nt	nt	nt	nt
Oral mucosa	12%	33%	nt	nt	nt	nt	nt	Mosaic
Fibroblasts	15%	45%	Mosaic	nt	Mosaic	nt	nt	nt
Hair roots	nt	nt	nt	41%	nt	nt	nt	nt
Genetic test	NGS	NGS	Sanger sequencing	Restriction digestion	Sanger sequencing	Sanger sequencing	Sanger sequencing	Sanger sequencing
Required diagnostic criteria for RTT								
Age of onset for a period of regression followed by recovery or stabilization	18 months	8-10 months	18 months	11 months	No regression	12 months	2 years	uk
Exclusion criteria for RTT	0/3	0/3	0/3	0/3	1/3	0/3	0/3	uk
Brain injury	No	No	No	No	No	No	No	uk
Neurometabolic disease, or severe infection	No	No	No	No	No	No	No	uk
Grossly abnormal development <6 months	No	No	No	No	Yes	No	No	uk
Main criteria for RTT	4/4	4/4	4/4	4/4	4/4	4/4	4/4	uk
Hand skills	Poor	Poor	Lost	Lost	None	Lost	Poor	uk
Language	Lost	None	Poor	None	None	None	Yes	uk
Gait	Ataxic	Poor	No	None	Ataxic	Lost	Ataxic	uk

(*Continues*)

TABLE 1 (Continued)

	Present cases		Published cases with new information		Previously published cases without new information			
	Case A	Case B	Psoni et al[9]	Topcu et al[8]	Pieras et al[10]	Kleefstra et al[11]	Clayton-Smith et al[6]	Armstrong et al[7a]
Gender	Male	Male	Male	Male	Male	Male	Male	Male
Stereotypic hand movements	Yes	Yes	Yes	Yes	Yes	Yes	Yes	uk
Supportive criteria for RTT	9/11	8/11	1/11	7/11	5/11	5/11	7/11	uk
Breathing disturbance	Yes	Yes	uk	No	No	uk	Yes	uk
Bruxism	Yes	Yes	uk	Yes	Yes	Yes	Yes	uk
Impaired sleep pattern	Yes	Yes	uk	uk	Yes	Yes	No	uk
Abnormal muscle tone	Yes	Yes	Yes	Yes	Yes	Yes	Yes	uk
Peripheral vasomotor disturbances	Yes	No	No	Yes	uk	uk	Yes	uk
Scoliosis/kyphosis	No	Yes	No	Yes	uk	uk	Yes	uk
Growth retardation	Yes	Yes	No	Yes	uk	uk	Yes	uk
Small cold hands and feet	Yes	Yes	No	Yes	uk	Yes	Yes	uk
Inappropriate laugh/ screaming spells	Yes	No	No	Yes	Yes	Yes	uk	uk
Diminished response to pain	Yes	Yes	uk	uk	uk	uk	uk	uk
Intense eye communication	No	No	No	uk	Yes	uk	uk	uk
RTT diagnosis (based on the Neul criteria 2010[2])	Classical RTT	Classical RTT	Classical RTT	Classical RTT	Classical RTT	Classical RTT	Classical RTT	Classical RTT, according to publication

Abbreviations: nt, not tested; RTT, Rett syndrome; uk, unknown.

[a] This case was published as having classical RTT without a detailed clinical description but as a patient with classical RTT and 8/9 of the symptoms according to the 2001 criteria.[2]

In conclusion, hand stereotypies are a clinical hallmark of RTT, and together with a regression period with loss of hand skills and speech, these unique symptoms can aid the clinician to consider RTT and apply the Neul clinical criteria in both males and females. This applies both to the clinical assessment before genetic testing and in cases where a *MECP2* variant is found without a prior suspicion of RTT.

Furthermore, the interpretation of NGS data should have a low threshold for variant calls in case-specific highly relevant genes, high coverage in critical gene regions, and examination of various tissues is relevant to confirm a low-frequent degree of mosaicism.

ACKNOWLEDGEMENTS

We thank the families for participating in this study. We also thank the laboratories at the Department of Clinical Genetics, Rigshospitalet and Odense University Hospital as well as the Murdoch Children's Research Institute for their laboratory expertise. This study was supported by "Dronning Louises Børnehospitals Fond." The research conducted at the Murdoch Children's Research Institute was supported by the Victorian Government's Operational Infrastructure Support Program.

ETHICAL APPROVAL

The project was approved by the Danish regional ethics committee and the regional Danish data protection agency.

CONFLICTS OF INTEREST

The authors declare no conflict of interests.

ORCID

Bitten Schönewolf-Greulich https://orcid.org/0000-0003-0088-0792

REFERENCES

1. Rett A. Über ein eigenartiges hirnatrophisches Syndrom bei Hyperammonämie im Kindersalter. *Wiener Medizinische Wochenschrift*. 1966;116(37):723-726.
2. Neul JL, Kaufmann WE, Glaze DG, et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol*. 2010;68(6):944-950.
3. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*. 1999;23(2):185-188.
4. Christodoulou J, Grimm A, Maher T, Bennetts B. RettBASE: the IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat*. 2003;21(5):466-472.
5. Villard L. MECP2 mutations in males. *J Med Genet*. 2007;44(7):417-423.
6. Clayton-Smith J, Watson P, Ramsden S, Black GC. Somatic mutation in MECP2 as a non-fatal neurodevelopmental disorder in males. *Lancet (London, England)*. 2000;356(9232):830-832.
7. Armstrong J, Pineda M, Aibar E, Gean E, Monros E. Classic Rett syndrome in a boy as a result of somatic mosaicism for a MECP2 mutation. *Ann Neurol*. 2001;50(5):692.
8. Topcu M, Akyerli C, Sayi A, et al. Somatic mosaicism for a MECP2 mutation associated with classic Rett syndrome in a boy. *Eur J Hum Genet*. 2002;10(1):77-81.
9. Psoni S, Sofocleous C, Traeger-Synodinos J, Kitsiou-Tzeli S, Kanavakis E, Fryssira-Kanioura H. Phenotypic and genotypic variability in four males with MECP2 gene sequence aberrations including a novel deletion. *Pediatr Res*. 2010;67(5):551-556.
10. Pieras JI, Munoz-Cabello B, Borrego S, et al. Somatic mosaicism for Y120X mutation in the MECP2 gene causes atypical Rett syndrome in a male. *Brain Dev*. 2011;33(7):608-611.
11. Kleefstra T, Yntema HG, Nillesen WM, et al. MECP2 analysis in mentally retarded patients: implications for routine DNA diagnostics. *Eur J Hum Genet*. 2004;12(1):24-28.
12. Leonard H, Silberstein J, Falk R, et al. Occurrence of Rett syndrome in boys. *J Child Neurol*. 2001;16(5):333-338.
13. Vorsanova SG, Yurov YB, Ulas VY, et al. Cytogenetic and molecular-cytogenetic studies of Rett syndrome (RTT): a retrospective analysis of a Russian cohort of RTT patients (the investigation of 57 girls and three boys). *Brain Dev*. 2001;23:S196-S201.
14. Stosser MB, Lindy AS, Butler E, et al. High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet Med*. 2018;20(4):403-410.
15. Schonewolf-Greulich B, Tejada MI, Stephens K, et al. The MECP2 variant c.925C>T (p.Arg309Trp) causes intellectual disability in both males and females without classic features of Rett syndrome. *Clin Genet*. 2016;89(6):733-738.
16. Nielsen JB, Henriksen KF, Hansen C, Silahtaroglu A, Schwartz M, Tommerup N. MECP2 mutations in Danish patients with Rett syndrome: high frequency of mutations but no consistent correlations with clinical severity or with the X chromosome inactivation pattern. *Eur J Hum Genet*. 2001;9:178-184.
17. Ravn K, Nielsen JB, Uldall P, Hansen FJ, Schwartz M. No correlation between phenotype and genotype in boys with a truncating *MECP2* mutation. *J Med Genet*. 2003;40(1):e5-e5, 55.
18. Bisgaard AM, Schonewolf-Greulich B, Ravn K, Ronde G. Is it possible to diagnose Rett syndrome before classical symptoms become obvious? Review of 24 Danish cases born between 2003 and 2012. *Eur J Paediatr Neurol*. 2015;19(6):679-687.
19. Acuna-Hidalgo R, Bo T, Kwint MP, et al. Post-zygotic point mutations are an Underrecognized source of De novo genomic variation. *Am J Hum Genet*. 2015;97(1):67-74.
20. Dou Y, Gold HD, Luquette LJ, Park PJ. Detecting somatic mutations in normal cells. *Trends Genet*. 2018;35(7):545-557.

SUPPORTING INFORMATION

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

How to cite this article: Schönewolf-Greulich B, Bisgaard A-M, Dunø M, et al. Mosaic *MECP2* variants in males with classical Rett syndrome features, including stereotypical hand movements. *Clin Genet*. 2019;95:403–408. https://doi.org/10.1111/cge.13473

Publication 2

Schonewolf-Greulich, B., A. M. Bisgaard, R. S. Moller, M. Duno, K. Brondum-Nielsen, **S. Kaur**, N. J. Van Bergen, S. Lunke, S. Eggers, C. Jespersgaard, J. Christodoulou and Z. Tumer (2019). Clinician's guide to genes associated with Rett-like phenotypes-Investigation of a Danish cohort and review of the literature. *Clin Genet, 95*(2), 221-230. doi:10.1111/cge.13153

Received: 4 September 2017 | Revised: 4 October 2017 | Accepted: 5 October 2017

DOI: 10.1111/cge.13153

WILEY CLINICAL GENETICS

REVIEW

Clinician's guide to genes associated with Rett-like phenotypes—Investigation of a Danish cohort and review of the literature

B. Schönewolf-Greulich[1,2] | A-M. Bisgaard[1] | R.S. Møller[3,4] | M. Dunø[5] | K. Brøndum-Nielsen[5] | S. Kaur[6,7] | N.J. Van Bergen[6,7] | S. Lunke[8] | S. Eggers[8] | C. Jespersgaard[2] | J. Christodoulou[6,7] | Z. Tümer[2]

[1]Center for Rett Syndrome, Kennedy Center, Department of Paediatrics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[2]Applied Human Molecular Genetics, Kennedy Center, Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[3]Danish Epilepsy Centre, Dianalund, Denmark

[4]Institute for Regional Health Services, University of Southern Denmark, Odense, Denmark

[5]Department of Clinical Genetics, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

[6]Neurodevelopmental Genomics Research Group, Murdoch Children's Research Institute, Melbourne, Australia

[7]Department of Paediatrics, Melbourne Medical School, University of Melbourne, Melbourne, Australia

[8]Translational Genomics Unit, Murdoch Children's Research Institute, Melbourne, Australia

Correspondence
Bitten Schönewolf-Greulich, Center for Rett Syndrome, Kennedy Center, Rigshospitalet, Gl. Landevej 7, 2600 Glostrup, Copenhagen, Denmark.
Email: bitten.schoenewolf-greulich.01@regionh.dk

Funding information
Fonden til Lægevidenskabens Fremme; Dr. Louises Børnehospitals Forskningsfond

The differential diagnostics in Rett syndrome has evolved with the development of next generation sequencing-based techniques and many patients have been diagnosed with other syndromes or variants in newly described genes where the associated phenotype(s) is yet to be fully explored. The term Rett-like refers to phenotypes with distinct overlapping features of Rett syndrome where the clinical criteria are not completely fulfilled. In this study we have combined a review of Rett-like disorders with data from a Danish cohort of 35 patients with Rett-like phenotypes emphasizing the diagnostic overlap with Pitt-Hopkins syndrome, Cornelia de Lange syndrome with *SMC1A* variants, and epileptic encephalopathies, for example, due to *STXBP1* variants. We also found a patient with a pathogenic variant in *KCNB1*, which has not been previously linked to a Rett-like phenotype. This study underlines the clinical and genetic heterogeneity of a Rett syndrome spectrum, and provides an overview of the Rett syndrome-related genes described to date, and hence serves as a guide for diagnosing patients with Rett-like phenotypes.

KEYWORDS

atypical Rett, *CDKL5*, *FOXG1*, *KCNB1*, *MECP2*, Rett-like, RTT, *SMC1A*

1 | INTRODUCTION

The neurodevelopmental disorder Rett syndrome (RTT) was first described in 1966 by Dr Andreas Rett, who reported 22 girls with loss of speech and hand use, and who developed hand stereotypies.[1] Apart from the classical/typical form (RTT) there are atypical forms of RTT (atypical RTT) with different clinical subgroups such as "preserved speech variant,"[2] the "early seizure variant"[3,4] and the "congenital variant,"[5] Today, the term Rett-like is used when there is an overlapping phenotype with typical or atypical RTT, but in whom the

John Christodoulou and Zeynep Tümer were co-last authors.

clinical criteria are not fulfilled. There are to date no formal published consensus criteria for a Rett-like diagnosis, and thus this description can be given to any patient with different combinations of the distinct features of RTT.

The genetic causes of typical RTT, atypical RTT and Rett-like disorders have been intensely studied. Typical RTT patients almost exclusively have pathogenic *MECP2* variants,[6] while atypical RTT and Rett-like patients show genetic heterogeneity. The most common genes involved apart from *MECP2* are: *CDKL5* (cyclin-dependent kinase-like 5), in which variants cause the early onset seizure disorder[7] and *FOXG1*, which is responsible for the congenital RTT variant.[8] Several additional new genes have been suggested to be associated with atypical RTT and a Rett-like disorder.[9–12] These studies have also revealed previously unrecognized clinical similarities of Rett spectrum phenotypes with other neurodevelopmental disorders. One of the unexpected findings has been the identification of *SMC1A* variants, which are traditionally linked to Cornelia de Lange syndrome (CdLS), in Rett-like patients.[13] Furthermore, several Rett-like patients have pathogenic variants in epileptic encephalopathy genes.[14] From a clinical perspective these findings enable an improved comprehension of the genetic landscape of atypical RTT and Rett-like disorders, potentially transforming the primary clinical diagnosis to a genetic one. From a research perspective, the genetic landscape is yet to completely unfold, and some of the genes associated with these phenotypes may be linked through similar functional pathways.

In this study, we present a genetic review of a Rett-like spectrum (typical RTT, atypical RTT and Rett-like disorder) and our findings in a Danish Rett-like cohort, including association of the epilepsy gene, *KCNB1*, as a new Rett-like gene. The purpose of this review is to provide a guide for molecular diagnosis of patients with Rett-like phenotypes.

2 | RETT SYNDROME

RTT is mainly seen in females and presents with only a few neurodevelopmental symptoms before the age of 6 to 18 months.[15] A period of regression will then occur, and abilities especially related to hand function and speech can deteriorate or be lost altogether.[16] Hand stereotypies are a main feature and the patients often have severely impaired functional use of their hands.[17] Another distinct clinical feature often present which can help to point to the diagnosis is the breathing abnormalities (including hyperventilation and/or breath holding episodes).[18] Patients all have intellectual disability (ID). Other features that often develop over time include progressive microcephaly[19] and epilepsy (occurring in approximately 68% at a mean age of 4.7 years).[20,21] Scoliosis is also very common.[16] Facial dysmorphism is not distinct, but few have subtle dysmorphic facial features which do not enable a clinical diagnosis.[22]

Today the international diagnostic criteria published in 2010[23] are used to assist in establishing the clinical diagnosis of RTT (listed in Table 1). RTT patients should have 4 main criteria (partial or complete loss of purposeful hand movements, partial or complete loss of spoken language, a gait abnormality, and stereotypic hand movements); while a diagnosis of atypical RTT is suggested if the patient has 2 or more of the main in addition to 5 or more of the supportive criteria.[23] Both forms are characterized by a period of developmental regression followed by recovery or stabilization. In addition, the presence of any of the following features excludes a diagnosis of typical or classic RTT: a brain injury, a neurometabolic disease or neurological infection, and abnormal psychomotor development with an onset before the age of 6 months.[23]

2.1 | The Rett syndrome genes

MECP2, encoding methyl CpG binding protein 2, is the major gene related to RTT and there are over 4600 *MECP2* variants (pathogenic, benign, or with unknown significance) described in the literature.[24] *MECP2* variants mainly cause RTT (including classical/typical and variant/atypical RTT), but may rarely also be associated with non-syndromic neurodevelopmental disorders including autism.[25]

Variants in *CDKL5* have been reported in about 500 patients[24] and the most notable clinical feature of patients with pathogenic mutations of this gene is early onset epilepsy. These patients have ID and progressive microcephaly.[26] Stereotypic hand movements, respiratory impairment with breath holding and hyperventilation, and severely impaired speech, hand function and gait may also be observed.[27,28] These features are suggestive of RTT, but normal development during the first 6 months is rare in these patients and they often have facial dysmorphic features.[29,30] Traditionally, the term Hanefeld or early onset epilepsy variant of RTT[4,26] has been used for this diagnosis, but recently the term CDKL5 disorder has been introduced and is considered as a separate clinical disorder from RTT.[31]

FOXG1 variants are less common (approximately 80 cases in total[29]) and neurodevelopmental impairment is already observable in early infancy. As *FOXG1* is located on chromosome 14, variants affect males as well as females.

This condition has therefore been called the congenital variant of RTT,[8] and recently the term FOXG1 syndrome has been introduced.[29] These patients exhibit profound motor dysfunction, and stereotypic hand movements and abnormal breathing can be present.[32] There is rarely any speech development and eye contact is poor.[32,33] Epilepsy is common although the age of onset can vary substantially.[34] They have progressive microcephaly[33,35] and MRI of the brain reveals poor myelinization and/or corpus callosum hypoplasia or agenesis.[32]

3 | MANY GENES, MANY DISORDERS

Apart from these 3 genes, there are at least 28 OMIM morbid genes (https://www.omim.org/) described in literature in which variants may cause gene disorders/syndromes with overlapping Rett-like phenotypes (Table 2).[9,10,14,36–38] Some of these genes are associated with well-known syndromes such as Pitt-Hopkins syndrome (*TCF4*, *CNTNAP2*, *NRXN1*), Cornelia de Lange syndrome (*SMC1A*), Phelan-McDermid syndrome (*SHANK3*), Kleefstra syndrome (*EHMT1*), Christianson type X-linked mental retardation syndrome (*SLC9A6*), Angelman syndrome (*UBE3A*), and Glass syndrome (*SATB2*). In addition,

TABLE 1 Summary of the clinical RTT features of the Danish Rett-like cases

Gene	*STXBP1*	*STXBP1*	*SCN2A*	*KCNB1*	*TCF4*	*SHANK3*	*SMC1A*
Gender	Female	Female	Female	Female	Female	Female	Female
Variation	c.1061G>A	c.230T>A	c.3955C>T	c.946G>A	c.1669dupC	c.2265 + 1G>A	c.3576delA
Inheritance	de novo	de novo	de novo	de novo	de novo	de novo	de novo
Head size	Microcephaly	Normal	Progressive microcephaly	Normal	Progressive microcephaly	Normal	Progressive microcephaly
Epilepsy	From birth to 6 mo	From 2 wk to 3 mo	From 17 mo	From 2 y	From 2.5 y	No	From 2.5 years
Other features					Dysmorphic features	Infections	
Diagnostic criteria for Rett syndrome							
Required diagnostic criteria							
Age for a period of regression followed by recovery or stabilization	1.5 y	None	17 mo	6 mo	None	5.5 y	2.5 y
Exclusion criteria for typical RTT	0/3	1/3	0/3	0/3	1/3	0/3	0/3
Brain injury	No	No	No	No	No	No	No
Neurometabolic disease or severe infection	No	No	No	No	No	No	No
Grossly abnormal development in first 6 mo	No	Yes	No	No	Yes	No	No
Main criteria	3/4	3/4	3/4	4/4	4/4	3/4	4/4
Hand skills	Partial loss	Impaired	Poor	Lost	Poor	Partial loss	Partial loss
Language	Lost	Never acquired	Never acquired	Never acquired	Never acquired	Partial loss	Lost
Gait	Partial loss	Impaired	Impaired	No gait	Impaired	Normal	No gait
Stereotypic hand movements	No	No	Stereotypies or automatisms	Hand wringing	Hand wringing	Hands to mouth	Hand wringing
Supportive criteria for atypical RTT	8/11	5/11	7/11	3/11	4/11	4/11	5/11
Breathing abnormality	Yes	No	Yes	Yes	Yes	No	Yes
Bruxism	Yes	Yes	Yes	No	No	Yes	No
Impaired sleep pattern	Yes	No	Yes	Yes	Yes	Yes	No
Abnormal muscle tone	Yes	Yes	Yes	Yes	Yes	No	Yes
Peripheral vasomotor disturbances	No	No	No	No	No	No	No
Scoliosis/kyphosis	Yes	Yes	No	No	No	No	Yes
Growth retardation	Yes	No	Yes	No	No	No	Yes
Small cold hands and feet	Yes	No	Yes	No	No	No	Yes
Inappropriate laughing/ screaming spells	No	No	Yes	No	Yes	Yes	No
Diminished response to pain	Yes	Yes	Yes	No	No	Yes	No
Intense eye communication	No	Yes	No	No	No	No	No
RTT diagnosis (based on the Neul criteria)	Atypical	Atypical	Atypical	Typical	RTT-like	RTT-like	Typical

The human phenotype ontology (HPO) terms corresponding to the symptoms, which are part of the Rett-syndrome diagnostic criteria are listed in Table S1.

TABLE 2 List of causative genes in Rett-like diagnoses or differential diagnoses and the phenotypical overlap with Rett syndrome

		Present in RTT										Not present in RTT			
		Required				4 main criteria				Other common symptoms		Exclusion criteria	Other symptoms		
		Developmental regression	Purposefull hand movements lost/absent	Speech severe deficit/loss	Gait abnormality	Stereotypic hand movements	Breathing abnormality	ID	Epilepsy	Microcephaly	CNS abnormality	Dysmorphic facial features	Other prominent symptoms	Disorders	OMIM#	Inheritance
RTT genes	*MECP2*	+	+	+	+	+	+	+	+	+	-	-		Rett Syndrome	312750	XLD
	CDKL5	+	+	+	+	+	+	+	+	+	-	+		Epileptic encephalopathy 2	300672	XLD
	FOXG1	+	+	+	+	+	-	+	+	+	+	-		Rett syndrome, congenital variant	613454	AD
Well-known syndrome genes	*SMC1A*	+	+	+	+	+	+	+	+	+	+	+		Cornelia de Lange syndrome 2	300590	XLD
	TCF4	-	+	+	+	+	+	+	+	+	+	+		Pitt-Hopkins syndrome	610954	AD
	CNTNAP2	+	+	+	+	+	+	+	+	+	+	+		Pitt-Hopkins like, syndrome 1	610042	AR
	NRXN1	-	+	+	+	+	+	+	+	+	+	+		Pitt-Hopkins like syndrome 2	614325	AR
	SHANK3	+	-	+	+	+	-	+	+	-	+	+		Phelan-McDermid syndrome	606232	AD
	SATB2	-	+	+	+	+	+	+	+	+	-	+	CP	Glass syndrome	612313	AD
	SLC9A6	+	-	+	+	-	-	+	+	+	+	+		Mental retardation, X-linked syndromic, Christianson type	300243	XLD
	EHMT1	+	-	+	+	-	-	+	+	+	+	-	CHD	Kleefstra syndrome	610253	AD
	UBE3A	-	-	+	+	+	-	+	+	+	+	+		Angelman syndrome	105830	AD
Genes of epilepsy	*STXBP1*	+	+	+	+	+	-	+	+	+	-	-		Epileptic encephalopathy 4	612164	AD
	SCN1A	+	+	+	+	+	-	+	+	+	-	+		Epileptic encephalopathy, early infantile, 6 (Dravet syndrome)	607208	AD
	SCN2A	+	+	+	+	+	-	+	+	+	-	-		Epileptic encephalopathy, 11	613721	AD
	SCN8A	+	-	+	+	+	-	+	+	+	-	-		Epileptic encephalopathy, 13	600702	AD
	GRIN2A	+	-	+	+	+	-	-	+	-	-	-		FESD (incl. Landau-Kleffner syndrome)	245570	AD
	GRIN2B	-	-	+	+	+	+	+	+	+	-	+		Epileptic encephalopathy, early infantile, 27	616139	AD
	HCN1	-	+	+	+	+	+	+	+	+	-	-		Epileptic encephalopathy, early infantile, 24	602780	AD
	SLC6A1	+	?	+	+	+	-	+	+	?	?	+		Myoclonic-atonic epilepsy	616421	AD
	KCNA2	+	+	+	+	-	-	+	+	+	+	+		Epileptic encephalopathy, early infantile, 32	616366	AD
	EEF1A2	-	-	+	+	+	+	+	+	+	-	+		Epileptic encephalopathy, early infantile 33	616409	AD
	KCNB1	+	+	+	+	+	-	+	+	+	-	-		Epileptic encephalopathy, early infantile, 26	616056	AD
Other genes	*WDR45*	+	+	+	+	+	-	+	+	+	+	+		Neurodegeneration with brain iron accumulation 5	300894	XLD
	ST3GAL5	+	+	+	+	+	-	+	+	+	-	+	SPD	Salt and pepper developmental regression syndrome	609056	AR
	IQSEC2	+	+	+	+	+	+	+	+	+	-	-		X-linked mental retardation, MRX1	309530	XLD
	MEF2C	-	+	+	+	+	-	+	+	-	+	+		*MEF2C* haploinsufficiency syndrome	613443	AD
	SLC35A2	-	+	+	+	+	-	+	+	+	+	+	RP	Congenital disorder of glycosylation type IIm	300896	XLD
	MFSD8	+	+	+	+	+	+	+	+	+	+	-	VL	Ceroid lipofuscinosis neuronal 7	610951	AR
	EIF2B2	+	+	+	+	+	-	+	+	+	+	-		Leukoencephalopathy with vanishing white matter	603895	AR
	ZNF238	+	-	+	+	+	+	+	-	+	+	+		Mental Retardation, autosomal dominant 22 (AD)	612337	AD

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; CHD, coronary heart disease; CP, cleft palate; FESD, epilepsy, focal, with speech disorder and with or without mental retardation; SPD, skin pigmentation disorder; RP, retinitis pigmentosa; VL, visual loss; XLD, X-linked dominant. Plus (+) is noted if the symptom has been described in one or more patients with pathogenic variant in the gene. The symptoms emphasized are the main clinical features according to the 2010 classification of clinical Rett and other very specific features of RTT. Gray color are clinical symptoms in common with RTT.

a substantial number of genes are epileptic encephalopathy genes (*STXBP1*, *SCN1A*, *SCN2A*, *SCN8A*, *GRIN2A*, *GRIN2B*, *HCN1*, *SLC6A1*, *KCNA2*, *EEF1A2* and *KCNB1*).

4 | DANISH COHORT OF RTT, ATYPICAL RTT AND RETT-LIKE PHENOTYPES

In Denmark, patients with clinical features suggestive of RTT are referred to the National Danish Centre for Rett syndrome and currently, we have registered 146 living patients. All these patients were initially investigated for the 3 known RTT genes and the sequence variant distribution is: *MECP2*: 99 cases, *CDKL5*: 10 cases, and *FOXG1*: 2 cases.

Among the patients, without deficiencies in 1 of these 3 genes (n = 35), pathogenic sequence variants were identified in *SCN2A* (n = 1), *STXBP1* (n = 2) and *TCF4* (n = 1); all genes previously linked to Rett-like phenotype (Table 2), using different technologies including single gene or whole exome sequencing, and a previously described epilepsy gene panel.[39] Notably, in one patient we identified a de novo variant in *KCNB1*, which was not previously linked to Rett syndrome spectrum. Patient 2 (*STXBP1*) and patient 3 (*SCN2A*) have previously been published[40,41] (Table 1).

Remaining patients were investigated using an in-house gene panel which included the genes described in Table 2. Through this approach we identified 2 patients with variants within *SMC1A* (n = 1) and *SHANK3* (n = 1) (Table 1). With these results, pathogenic variants of a total of 23 different have been found in RTT/atypical RTT/Rett-like patients (Figure 1).

The project was approved by the Danish regional ethics committee, the regional data protection agency and informed consent was obtained from all participating patients.

5 | CLINICAL OR GENETIC DIAGNOSIS?

The description of variants within several different genes in patients with overlapping phenotypes creates practical challenges with diagnosis in clinical practice. Next generation sequencing-based methodologies (gene panels or whole exome/genome sequencing) have overcome some difficulties and diagnosis tends to move from a sole clinical to a genetic confirmation. However, understanding the clinical

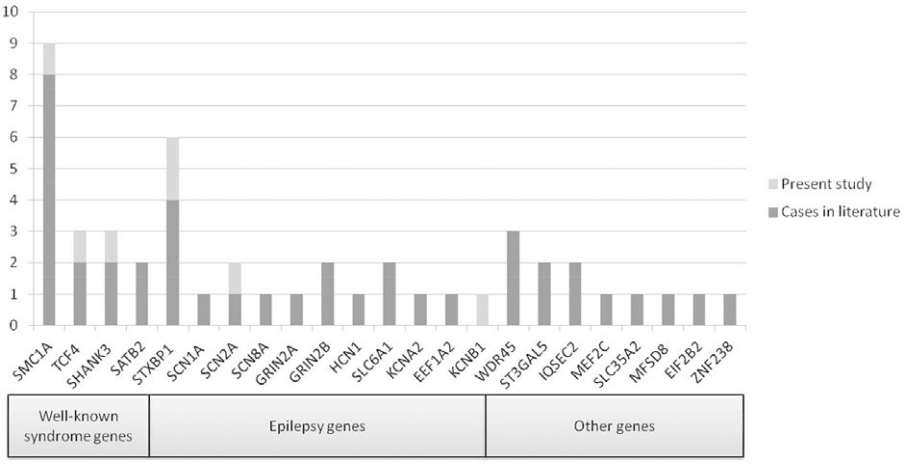


FIGURE 1 Twenty-three OMIM genes found in Rett-like phenotypes in literature (dark gray) and in the Danish database (light gray)

picture of different disorders is crucial for selection of the most relevant genes to be investigated and for assessing the clinical significance of genomic variants identified in these genes. In the following section we will compare different disorders and causative genes in comparison with RTT features (Table 2).

5.1 | Pitt-Hopkins and Pitt-Hopkins-like syndromes

Pitt-Hopkins syndrome was first described in 1978 in patients with distinct facial dysmorphism and overbreathing.[42] Neurodevelopment and speech are severely impaired, and several patients have progressive microcephaly and seizures.[28,43] Gait is often ataxic. Fetal pads on fingers and toes, and stereotypic hand movements such as hand clapping are important features. *TCF4* is the main causative gene for the classical Pitt-Hopkins phenotype;[44] variants in this gene have also been reported in 2 patients with a Rett-like phenotype.[10,12] In the Danish cohort, a single patient suspected of having RTT due to hyperventilation and hand stereotypies has a pathogenic *TCF4* variant (Table 1).

Bi-allelic loss-of-function variants in *CNTNAP2* or *NRXN1* can also result in a Pitt-Hopkins-like phenotype.[45] These patients also have severe ID and in some regression or stagnation in development.[45,46] Other prominent symptoms include epileptic seizures, a speech deficit with loss of verbal skills and hyperventilation; and some of the patients are described as being non-dysmorphic.[46] Even though variants of these genes are not yet found in Rett-like patients, they should be included in the genomic analyses of this patient group due the phenotypic overlap.

5.2 | Cornelia de Lange syndrome

CdLS is characterized by ID, a recognizable pattern of dysmorphic facial features (synophrys, thick-arched eyebrows, long eye lashes and downslanting palpebral fissures), hirsuitism, short neck, and possible presence of major malformations, such as limb defects.[47] There are 5 known genes associated with CdLS. Mutations in *SMC1A* are the second most common cause of CdLS, and commonly associated

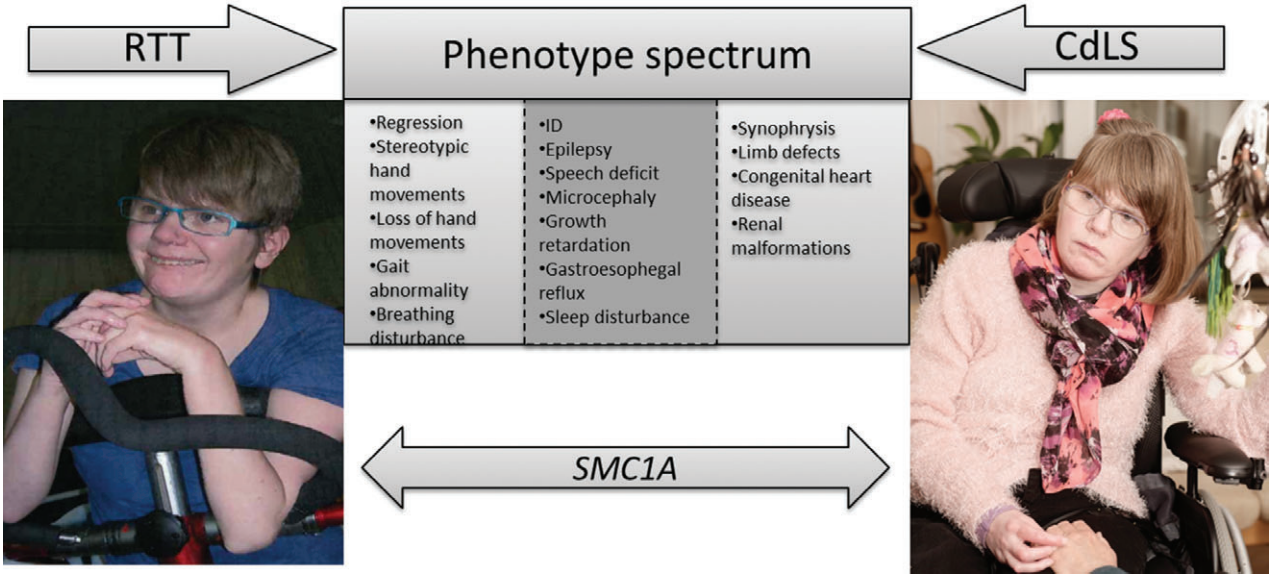


FIGURE 2 Clinical spectrum of *SMC1A* spanning from Rett syndrome to Cornelia de lange syndrome (CdLS). The photographs belong to the same patient. On the left photograph her hand wringing can be seen, but the eyebrows are covered. On the right side the synophrysis is slightly observable, but her hand movements cannot be seen as her hands are being held. In this respect she has the 2 very characteristic symptoms of both syndromes simultaneously. Photography on the right courtesy of Sofie Amalie Kloughart

with seizures, but limb defects are rare.[47] The phenotype associated with *SMC1A* variants may in some include Rett-like features[47,48] (Figure 2). They may have hand stereotypies, apnea, microcephaly, and some have a period of regression. *SMC1A* variants have been described in several patients with a Rett-like disorder,[11,49] and we identified a single patient with atypical RTT in the Danish cohort (Figure 1). The phenotypic spectrum of *SMC1A* spans from CdLS to Rett-like phenotypes, and *SMC1A* can now be considered as the fourth most common gene mutated in patients with a clinical Rett syndrome spectrum (Figure 1).

5.3 | Phelan-McDermid syndrome and *SHANK3*

Phelan-McDermid syndrome patients have various degrees of ID and speech deficits, epilepsy, hypotonia, structural brain and/or renal abnormalities and prevalent dysmorphic features, including a pointed chin and dysplastic toe nails.[50,51] The syndrome is caused by microdeletions encompassing *SHANK3* at 22q13 or a pathogenic variant in this gene, which is also associated with autism.[52] *SHANK3* variants have been reported in 2 patients with RTT[53] and a Rett-like phenotype,[12] respectively. In the Danish cohort a *SHANK3* variant was identified in a single patient, who fulfills some of the main criteria of RTT, but due to her normal gait, not categorized as typical RTT. On the other hand, she has classical autistic behavior compatible with a *SHANK3* variant.[52]

5.4 | Christianson type X-linked mental retardation

Christianson type X-linked mental retardation syndrome was first described in 1999 in a large family with males affected with profound ID, absence of speech and epilepsy.[54] Since then, several patients with common features (such as ataxic gait, hyperkinetic movements, a happy demeanor and poor fine motor skills) have been reported.[55,56] Epilepsy is reported in approximately 80% of cases, and some had a period of regression.[55,56] Because of this combination of clinical features, the syndrome has been considered a differential diagnosis of Angelman syndrome.[57] Christianson type X-linked mental retardation is caused by pathogenic variants in *SLC9A6* and has also been identified in 2 Rett-like patients.[11,12]

5.5 | *SATB2*-associated syndrome (Glass syndrome)

Pathogenic variants in *SATB2* have been identified in patients with neurodevelopmental disability, speech deficits and non-specific dysmorphic features, which together have been named as Glass syndrome or "*SATB2*-associated syndrome." Many patients have cleft palate or dental anomalies, but seizures were not prominent and they do not exhibit neurodevelopmental regression.[58,59] *SATB2* variants have also been described in 2 patients with a Rett-like phenotype, both of whom had hand stereotypies, impaired hand function and ID, but had no period of regression or seizures.[37] *SATB2*-associated syndrome should be considered when evaluating Rett-like patients, especially if they have oral anomalies.

5.6 | Kleefstra syndrome

The key features of Kleefstra syndrome are ID, hypotonia, dysmorphic features and cardiac defects.[36,60] However, some characteristic symptoms overlapping with RTT, such as speech loss and impaired walking, may be present.[61] Additional non-specific RTT features, including microcephaly, a sleep disorder and teeth grinding, have been described numerously.[60] Some patients have stereotypic movements and respiratory abnormalities, but are different to those observed in RTT patients.[36,61] Kleefstra syndrome is caused by 9q34.3 microdeletions including *EHMT1* or sequence variants of this gene.[60] *EHMT1* variants have not yet been reported in patients with a Rett-like phenotype, but have great phenotypic overlap and has been linked to the Rett-like spectrum.[36]

5.7 | Angelman syndrome

Angelman syndrome (AS) and RTT have often been described as potential differential diagnoses of each other, because they both present with severe ID, microcephaly, speech deficits and abnormal hand movements.[36] However, there are clinically identifiable differences between AS and RTT: AS patients have facial dysmorphism and they do not have hand wringing, but hand flapping. The EEG is almost always abnormal and often with a classical AS pattern.[35] AS can be caused by several genetic/epigenetic mechanisms, including variations in *UBE3A*,[62] variants of which have not yet been reported in Rett syndrome patients.

5.8 | Epilepsy

More than 500 genes have been associated with epilepsy, including genes causing encephalopathy, and pathogenic variants of several of the latter have also been described in patients with Rett-like features (Table 2, Figure 1). In Figure 3, we have compiled the age of onset of seizures for all the genes that have been associated with RTT or Rett-like clinical spectrum, as the type and age of onset of epilepsy can be an important indicator of which group of genes is probably to be causative. The epileptic encephalopathies may overlap with Rett-like phenotypes, especially where there is a period of regression with the onset of epilepsy, together with progressive microcephaly, and intellectual, speech and motor deficits.[14] Epileptic encephalopathies are defined as conditions where the progressive disturbance in cerebral function is thought to be the result of the epileptiform abnormalities.[63] However, in some of the so-called epileptic encephalopathies it is probably that the neurodevelopmental co-morbidities are primary, because neurodevelopmental abnormalities can occur before the onset of seizures, and in fact some patients never experience any seizures. This has been the case with *STXBP1* variants, for instance, where patients have moderate to severe ID, speech deficits, ataxia and tremor[64–66]; epilepsy is described in 95% of cases with a median age of 6 weeks but ranging from 1 day to 12 years.[65] Six patients with a Rett-like phenotype, including 2 Danish patients, have been found to have *STXBP1* variants[10,11,14,67] (Figure 1).

Variants in *SCN1A*, *SCN2A* and *SCN8A* lead to early onset of seizures and variants in these genes have been described in single cases with Rett-like features. The development of epilepsy is variable in

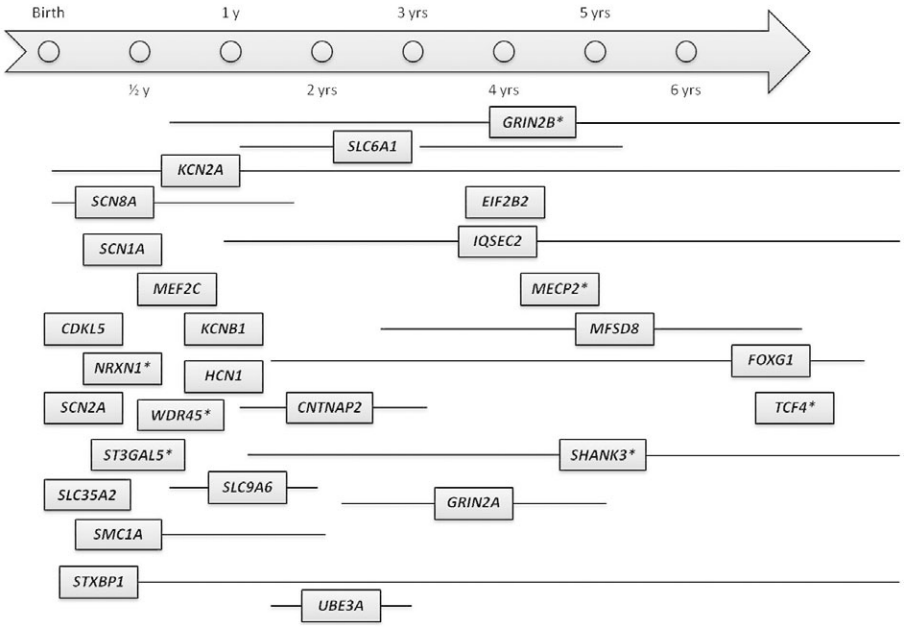


FIGURE 3 Average time for first epileptic seizure.* indicates that >5% of patients with variations in this gene have been reported without epilepsy

these cases, but in all of them a period of regression has followed the epileptic seizures.[12,14,68] In addition, patients may have ataxia, encephalopathy, various degrees of ID and hand stereotypies, and so these genes should be considered in Rett-like patients with early onset epilepsy. The severity of epilepsy is variable in patients with *SCN1A*, and an *SCN1A* variant was found in 1 patient with a Rett-like phenotype.[12] In the Danish cohort we identified a single patient with an *SCN2A* variant. She was referred with the suspicion of RTT, but her clinical history was more suggestive of a primary epileptic encephalopathy disorder, because the regression was related to the onset of intractable epilepsy.

GRIN2A and *GRIN2B* are 2 genes which cause ID and childhood onset epilepsy.[69] Variants in *GRIN2A* cause a clinical heterogeneous disorder, and have been reported in 20% of patients with Landau-Kleffner syndrome, continuous spike and waves during slow-wave sleep syndrome and electroclinically atypical rolandic epilepsy.[70] *GRIN2B* is currently considered an epileptic encephalopathy gene,[71,72] but variants have been reported in patients both with and without epilepsy. Evaluation of larger cohorts is needed to better define the phenotype associated with variants in this gene. Notably, variations in *GRIN2A* and *GRIN2B* have been reported in 3 Rett-like patients, all of whom had hand stereotypies, whereas the patient with a *GRIN2B* variant also had a breathing disturbance.[11,12]

Variants of *HCN1* and *SLC6A1* have been reported only in few patients in epileptic encephalopathy cohorts[73,74]; and *EEF1A2* variants were detected in a few patients with ID, dysmorphic facial features epilepsy with a variable age of onset.[75] Variants in these genes are described in single patients with a Rett-like phenotype, including stereotypic hand movements and breathing abnormalities as distinct features of RTT.[11,12]

KCNA2 variants are phenotypically linked to infantile epilepsy and episodic ataxia,[9] but also to hereditary spastic paraplegia.[76] The only Rett-like phenotype patient with a *KCNA2* variant had an onset of regression a few months before the onset of intractable epilepsy. She also had loss of hand skills, hand stereotypies, microcephaly and a speech deficit.[9]

KCNB1 has previously been described as a gene related to epileptic encephalopathy and ID, and sequence variants have only been described in a few patients, one of whom had handwringing.[77,78] *KCNB1* variants have not been previously described in Rett-like patients, even though they may be associated with clinical features overlapping with those of RTT. In the Danish cohort we identified a single case with pathogenic *KCNB1* variant. The patient displayed many RTT features, including neurodevelopmental regression, loss of hand function, hand stereotypies and a speech deficit, and the onset of cognitive and motor decline was prior to her seizures. We suggest that this finding warrants the inclusion of *KCBN1* in the expanding group of Rett-like genes.

5.9 | *WDR45* and other genes

WDR45 variants were first linked to a neurodegenerative disorder with iron deposits, and have been reported in 4 patients with an initial diagnosis of RTT or a Rett-like disorder, mainly because of cognitive and motor decline and hand stereotypies. Patients with *WDR45* variants can be interpreted as Rett-like in the early stages of development before the brain MRI scans displaying the iron deposits typical of WDR45 disorder become obvious.[11,79]

Variants of all the other genes listed in Figure 1,[10,37,80–82] except for *WDR45*,[38,79,83] have been observed in only 1 or 2 patients, and these genes are described in more detail in Appendix S1 (Supporting information).

6 | CLINICAL SIGNATURES OF RETT-LIKE PHENOTYPES

In a cohort of 35 Danish patients with Rett-like features but without *MECP2*, *CDKL5* or *FOXG1* variants, 6 patients revealed pathogenic changes in 5 other genes (*SMC1A*, *TCF4*, *SHANK3*, *STXBP1* and *SCN2A*), previously linked to a Rett-like phenotype. Furthermore, we

identified a *KCNB1* variant, expanding the range of "Rett spectrum genes."

These findings, together with the published reports, underscore that RTT and Rett-like phenotypes represent a clinically heterogeneous group of disorders, which may be designated as being part of a Rett syndrome spectrum (Figure 1). The underlying genetic defects are also quite heterogeneous, with pathogenic variants reported in a number of genes, but where variants of each gene are detected in only a few patients, as outlined above. However, a number of recurrences have been reported in a number of genes such as *SMC1A*, *TCF4*, *SHANK3* and *WDR45*, as well as a group of epileptic encephalopathy genes, all of which should be considered in the molecular diagnosis of patients with Rett-like phenotypes.

The reports of Andreas Rett and Bengt Hagberg and international colleagues, who undertook thorough initial evaluations of patients with RTT, unique co-morbidities have been identified.[16] These co-morbidities distinguish patients with RTT from those who have other neurodevelopmental disorders. Some of the recent genes associated with Rett-like phenotypes were detected through a genotype first approach as a consequence of large-scale genetic screening technologies, but not based on the clinical phenotype of the patients. However, it is possible to describe co-morbidities when patients are extensively phenotyped and compared with an adequate cohort of cases with mutations in the same gene. Descriptions of genes in a specific group of patients with a single major clinical feature like epilepsy may not emphasize other features or co-morbidities. Therefore, the full spectrum of the clinical picture will only be revealed when larger cohorts of patients with pathogenic variants of the same gene are described. This can lead to the "lumper" and "splitter" conundrum, where, for instance, newly identified genes associated with a Rett-like phenotype and severe epilepsy may be perceived as primary epileptic encephalopathies, or may be perhaps cast as separate syndromes based on their spectrum of co-morbidities. This has been discussed especially with regards to *STXBP1* variants[65,84] that have been recurrently reported in Rett-like patients (Figure 1). We also suggest a new "Rett-like gene," namely *KCNB1*, where several patients harboring mutations in this gene have handwringing and regression.

In the Danish cohort we found that using the clinical criteria described by Neul et al, the diagnosis of RTT is often very clear and pinpoints to a phenotype associated with *MECP2* variants with great accuracy.[23] However, it can be argued if the atypical RTT criteria really apply to patients with variants in *CDKL5* and *FOXG1*, because these patients often have early neurodevelopmental symptoms and/or early onset epilepsy. We find that when diagnosing Rett-like patients it is helpful to establish the age of onset of epilepsy as a pointer to the correct molecular diagnosis (Figure 3). In the Danish cohort, a female patient was molecularly diagnosed, at the age of 52, with a *CDKL5* variant by direct sequencing, which was performed because clinical examination and a thorough medical history revealed a period of severe epilepsy in early infancy. Thus, it should always be remembered that the phenotype may evolve with time. However, one must be cautious about applying hard and definite rules. For instance, for *FOXG1* the mean age of onset of seizures has been reported as 6.8 years, but seizures can occur as early as 6 months of age, and not all of the patients have seizures (87%).[34] Thus, a variant in this gene may be considered if there are other features suggestive of a *FOXG1* defect, regardless age of onset of seizures (Figure 3).

When evaluating a patient who does not fulfill the clinical RTT criteria but has overlapping features, some key features should be taken into consideration in order to establish the diagnosis: psychomotor development should be carefully assessed for the first year of life to evaluate the age of onset of regression and the age of onset of epilepsy. Furthermore, careful characterization of hand movements and breathing pattern are invaluable. Because of the clinical and genetic heterogeneity, the diagnostic approach for patients with a Rett-like phenotype is shifting to a genotype first approach using gene panels, whole exome or genome sequencing. Nevertheless the clinical evaluation with identification of distinct clinical features is still highly crucial to assess the validity of causation of the putative sequence variants. Furthermore, as more patients with variants in the same genes are identified, clinical subgroups can be outlined.

7 | PERSPECTIVES

From a clinical perspective an overview of the genes involved in typical RTT, atypical RTT and Rett-like phenotypes enables a better understanding of the genetic landscape of the RTT spectrum. From a research perspective, the genes associated with Rett-like phenotypes need to be elucidated further in order to gain a more complete understanding of the genetic and functional networks involved. Apart from the known genes, variants in a number of other genes with yet unknown functions have been found through screening of large Rett-like phenotype cohorts.[10,12] Some of these may represent new candidate genes and further investigations of future cohorts and functional studies are necessary to establish their potential role in this disease pathogenesis. Furthermore, elucidation of the functional pathways involved could pave the way for the development of future targeted therapies.

ACKNOWLEDGEMENTS

We thank the patients and their families for participating in the project. The research has been funded by "Dr. Louises Børnehospitals Fond" and "Fonden til Lægevidenskabens Fremme."

Conflict of interest

The authors declare no conflict of interests.

ORCID

B. Schönewolf-Greulich https://orcid.org/0000-0003-0088-0792

REFERENCES

1. Rett A. On a remarkable syndrome of cerebral atrophy associated with hyperammonaemia in childhood. *Wien Med Wochenschr*. 2016;166(11–12):322–324.
2. Zappella M. The Rett girls with preserved speech. *Brain Dev*. 1992;14(2):98–101.
3. Hagberg BA, Skjeldal OH. Rett variants: a suggested model for inclusion criteria. *Pediatr Neurol*. 1994;11(1):5–11.

4. Hanefeld F. The clinical pattern of the Rett syndrome. *Brain Dev*. 1985;7(3):320–325.
5. Rolando S. Rett syndrome: report of eight cases. *Brain Dev*. 1985;7(3):290–296.
6. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*. 1999;23(2):185–188.
7. Weaving LS, Christodoulou J, Williamson SL, et al. Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. *Am J Hum Genet*. 2004;75(6):1079–1093.
8. Ariani F, Hayek G, Rondinella D, et al. FOXG1 is responsible for the congenital variant of Rett syndrome. *Am J Hum Genet*. 2008;83(1):89–93.
9. Allou L, Julia S, Amsallem D, El Chehadeh S, Lambert L, Thevenon J, Duffourd Y, Saunier A, Bouquet P, Pere S, Moustaïne A, Ruaud L, Roth V, Jonveaux P, Philippe C Rett-like phenotypes: expanding the genetic heterogeneity to the KCNA2 gene and first familial case of CDKL5-related disease. *Clin Genet* 2016, 91, 431-440.
10. Lopes F, Barbosa M, Ameur A, et al. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet*. 2016;53(3):190–199.
11. Sajan SA, Jhangiani SN, Muzny DM, et al. Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet Med*. 2016;19:13–19.
12. Lucariello M, Vidal E, Vidal S, et al. Whole exome sequencing of Rett syndrome-like patients reveals the mutational diversity of the clinical phenotype. *Hum Genet*. 2016;135(12):1343–1354.
13. Lebrun N, Lebon S, Jeannet PY, Jacquemont S, Billuart P, Bienvenu T. Early-onset encephalopathy with epilepsy associated with a novel splice site mutation in SMC1A. *Am J Med Genet A*. 2015;167A(12):3076–3081.
14. Olson HE, Tambunan D, LaCoursiere C, et al. Mutations in epilepsy and intellectual disability genes in patients with features of Rett syndrome. *Am J Med Genet A*. 2015;167A(9):2017–2025.
15. Bisgaard AM, Schonewolf-Greulich B, Ravn K, Ronde G. Is it possible to diagnose Rett syndrome before classical symptoms become obvious? Review of 24 Danish cases born between 2003 and 2012. *Eur J Paediatr Neurol*. 2015;19(6):679–687.
16. Hagberg B. Clinical manifestations and stages of Rett syndrome. *Ment Retard Dev Disabil Res Rev*. 2002;8(2):61–65.
17. Downs J, Bebbington A, Jacoby P, et al. Level of purposeful hand function as a marker of clinical severity in Rett syndrome. *Dev Med Child Neurol*. 2010;52(9):817–823.
18. Pini G, Bigoni S, Congiu L, et al. Rett syndrome: a wide clinical and autonomic picture. *Orphanet J Rare Dis*. 2016;11(1):132.
19. Schultz RJ, Glaze DG, Motil KJ, et al. The pattern of growth failure in Rett syndrome. *Am J Dis Child*. 1993;147(6):633–637.
20. Lim Z, Downs J, Wong K, Ellaway C, Leonard H. Expanding the clinical picture of the MECP2 duplication syndrome. *Clin Genet*. 2016;91:557–563.
21. Nissenkorn A, Levy-Drummer RS, Bondi O, et al. Epilepsy in Rett syndrome—lessons from the Rett networked database. *Epilepsia*. 2015;56(4):569–576.
22. Allanson JE, Hennekam RC, Moog U, Smeets EE. Rett syndrome: a study of the face. *Am J Med Genet A*. 2011;155A(7):1563–1567.
23. Neul JL, Kaufmann WE, Glaze DG, et al. Rett syndrome: revised diagnostic criteria and nomenclature. *Ann Neurol*. 2010;68(6):944–950.
24. Krishnaraj R, Ho G, Christodoulou J. RettBASE: Rett syndrome database update. *Hum Mutat*. 2017;38:922–931.
25. Schonewolf-Greulich B, Tejada MI, Stephens K, et al. The MECP2 variant c.925C>T (p.Arg309Trp) causes intellectual disability in both males and females without classic features of Rett syndrome. *Clin Genet*. 2016;89(6):733–738.
26. Evans JC, Archer HL, Colley JP, et al. Early onset seizures and Rett-like features associated with mutations in CDKL5. *Eur J Hum Genet*. 2005;13(10):1113–1120.
27. Fehr S, Downs J, Ho G, et al. Functional abilities in children and adults with the CDKL5 disorder. *Am J Med Genet A*. 2016;170(11):2860–2869.
28. Marangi G, Zollino M. Pitt-Hopkins syndrome and differential diagnosis: a molecular and clinical challenge. *J Pediatr Genet*. 2015;4(3):168–176.
29. Mitter D, Pringsheim M, Kaulisch M, et al. FOXG1 syndrome: genotype-phenotype association in 83 patients with FOXG1 variants. *Genet Med*. 2017. Epub ahead of print.
30. Archer HL, Evans J, Edwards S, et al. CDKL5 mutations cause infantile spasms, early onset seizures, and severe mental retardation in female patients. *J Med Genet*. 2006;43(9):729–734.
31. Fehr S, Wilson M, Downs J, et al. The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *Eur J Hum Genet*. 2013;21(3):266–273.
32. Kortum F, Das S, Flindt M, et al. The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. *J Med Genet*. 2011;48(6):396–406.
33. Mencarelli MA, Spanhol-Rosseto A, Artuso R, et al. Novel FOXG1 mutations associated with the congenital variant of Rett syndrome. *J Med Genet*. 2010;47(1):49–53.
34. Seltzer LE, Ma M, Ahmed S, et al. Epilepsy and outcome in FOXG1-related disorders. *Epilepsia*. 2014;55(8):1292–1300.
35. Thibert RL, Larson AM, Hsieh DT, Raby AR, Thiele EA. Neurologic manifestations of Angelman syndrome. *Pediatr Neurol*. 2013;48(4):271–279.
36. Willemsen MH, Rensen JH, van Schrojenstein-Lantman de Valk HM, Hamel BC, Kleefstra T. Adult phenotypes in Angelman- and Rett-like syndromes. *Mol Syndromol*. 2012;2(3–5):217–234.
37. Lee JS, Yoo Y, Lim BC, Kim KJ, Choi M, Chae JH. SATB2-associated syndrome presenting with Rett-like phenotypes. *Clin Genet*. 2016;89(6):728–732.
38. Okamoto N, Ikeda T, Hasegawa T, et al. Early manifestations of BPAN in a pediatric patient. *Am J Med Genet A*. 2014;164A(12):3095–3099.
39. Moller RS, Larsen LH, Johannesen KM, et al. Gene panel testing in epileptic encephalopathies and familial epilepsies. *Mol Syndromol*. 2016;7(4):210–219.
40. Weckhuysen S, Holmgren P, Hendrickx R, et al. Reduction of seizure frequency after epilepsy surgery in a patient with STXBP1 encephalopathy and clinical description of six novel mutation carriers. *Epilepsia*. 2013;54(5):e74–e80.
41. Wolff M, Johannesen KM, Hedrich UB, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain*. 2017;140:1316–1336.
42. de Winter CF, Baas M, Bijlsma EK, van Heukelingen J, Routledge S, Hennekam RC. Phenotype and natural history in 101 individuals with Pitt-Hopkins syndrome through an internet questionnaire system. *Orphanet J Rare Dis*. 2016;11:37.
43. Peippo M, Ignatius J. Pitt-Hopkins Syndrome. *Mol Syndromol*. 2012;2(3–5):171–180.
44. Amiel J, Rio M, de Pontual L, et al. Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am J Hum Genet*. 2007;80(5):988–993.
45. Zweier C, de Jong EK, Zweier M, et al. CNTNAP2 and NRXN1 are mutated in autosomal-recessive Pitt-Hopkins-like mental retardation and determine the level of a common synaptic protein in drosophila. *Am J Hum Genet*. 2009;85(5):655–666.
46. Smogavec M, Cleall A, Hoyer J, et al. Eight further individuals with intellectual disability and epilepsy carrying bi-allelic CNTNAP2 aberrations allow delineation of the mutational and phenotypic spectrum. *J Med Genet*. 2016;53(12):820–827.
47. Boyle MI, Jespersgaard C, Brondum-Nielsen K, Bisgaard AM, Tumer Z. Cornelia de Lange syndrome. *Clin Genet*. 2015;88(1):1–12.
48. Huisman S, Mulder PA, Redeker E, et al. Phenotypes and genotypes in individuals with SMC1A variants. *Am J Med Genet A*. 2017;173:2108–2125.
49. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014;511(7509):344–347.
50. Sarasua SM, Dwivedi A, Boccuto L, et al. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). *J Med Genet*. 2011;48(11):761–766.
51. Kolevzon A, Angarita B, Bush L, et al. Phelan-McDermid syndrome: a review of the literature and practice parameters for medical assessment and monitoring. *J Neurodev Disord*. 2014;6(1):39.
52. Boccuto L, Lauri M, Sarasua SM, et al. Prevalence of SHANK3 variants in patients with different subtypes of autism spectrum disorders. *Eur J Hum Genet*. 2013;21(3):310–316.

53. Hara M, Ohba C, Yamashita Y, Saitsu H, Matsumoto N, Matsuishi T. De novo SHANK3 mutation causes Rett syndrome-like phenotype in a female patient. *Am J Med Genet A*. 2015;167(7):1593–1596.
54. Christianson AL, Stevenson RE, van der Meyden CH, et al. X linked severe mental retardation, craniofacial dysmorphology, epilepsy, ophthalmoplegia, and cerebellar atrophy in a large South African kindred is localised to Xq24-q27. *J Med Genet*. 1999;36(10):759–766.
55. Gilfillan GD, Selmer KK, Roxrud I, et al. SLC9A6 mutations cause X-linked mental retardation, microcephaly, epilepsy, and ataxia, a phenotype mimicking Angelman syndrome. *Am J Hum Genet*. 2008;82(4):1003–1010.
56. Schroer RJ, Holden KR, Tarpey PS, et al. Natural history of Christianson syndrome. *Am J Med Genet A*. 2010;152A(11):2775–2783.
57. Tan WH, Bird LM, Thibert RL, Williams CA. If not Angelman, what is it? A review of Angelman-like syndromes. *Am J Med Genet A*. 2014;164A(4):975–992.
58. Bengani H, Handley M, Alvi M, et al. Clinical and molecular consequences of disease-associated de novo mutations in SATB2. *Genet Med*. 2017;19:900–908.
59. Zarate YA, Kalsner L, Basinger A, et al. Genotype and phenotype in 12 additional individuals with SATB2-associated syndrome. *Clin Genet*. 2017;92:423–429.
60. Kleefstra T, Brunner HG, Amiel J, et al. Loss-of-function mutations in euchromatin histone methyl transferase 1 (EHMT1) cause the 9q34 subtelomeric deletion syndrome. *Am J Hum Genet*. 2006;79(2):370–377.
61. Verhoeven WM, Egger JI, Vermeulen K, van de Warrenburg BP, Kleefstra T. Kleefstra syndrome in three adult patients: further delineation of the behavioral and neurological phenotype shows aspects of a neurodegenerative course. *Am J Med Genet A*. 2011;155A(10):2409–2415.
62. Kishino T, Lalande M, Wagstaff J. UBE3A/E6-AP mutations cause Angelman syndrome. *Nat Genet*. 1997;15(1):70–73.
63. Engel J Jr. A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: report of the ILAE task force on classification and terminology. *Epilepsia*. 2001;42(6):796–803.
64. Gburek-Augustat J, Beck-Woedl S, Tzschach A, Bauer P, Schoening M, Riess A. Epilepsy is not a mandatory feature of STXBP1 associated ataxia-tremor-retardation syndrome. *Eur J Paediatr Neurol*. 2016;20(4):661–665.
65. Stamberger H, Nikanorova M, Willemsen MH, et al. STXBP1 encephalopathy: a neurodevelopmental disorder including epilepsy. *Neurology*. 2016;86(10):954–962.
66. Helbig I, Tayoun AA. Understanding genotypes and phenotypes in epileptic encephalopathies. *Mol Syndromol*. 2016;7(4):172–181.
67. Romaniello R, Saettini F, Panzeri E, Arrigoni F, Bassi MT, Borgatti R. A de-novo STXBP1 gene mutation in a patient showing the Rett syndrome phenotype. *Neuroreport*. 2015;26(5):254–257.
68. Baasch AL, Huning I, Gilissen C, et al. Exome sequencing identifies a de novo SCN2A mutation in a patient with intractable seizures, severe intellectual disability, optic atrophy, muscular hypotonia, and brain abnormalities. *Epilepsia*. 2014;55(4):e25–e29.
69. von Stülpnagel C, Ensslen M, Moller RS, et al. Epilepsy in patients with GRIN2A alterations: genetics, neurodevelopment, epileptic phenotype and response to anticonvulsive drugs. *Eur J Paediatr Neurol*. 2017;21(3):530–541.
70. Lesca G, Rudolf G, Bruneau N, et al. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat Genet*. 2013;45(9):1061–1066.
71. Endele S, Rosenberger G, Geider K, et al. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet*. 2010;42(11):1021–1026.
72. Lemke JR, Lal D, Reinthaler EM, et al. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet*. 2013;45(9):1067–1072.
73. Carvill GL, McMahon JM, Schneider A, et al. Mutations in the GABA transporter SLC6A1 cause epilepsy with myoclonic-atonic seizures. *Am J Hum Genet*. 2015;96(5):808–815.
74. Nava C, Dalle C, Rastetter A, et al. De novo mutations in HCN1 cause early infantile epileptic encephalopathy. *Nat Genet*. 2014;46(6):640–645.
75. Nakajima J, Okamoto N, Tohyama J, et al. De novo EEF1A2 mutations in patients with characteristic facial features, intellectual disability, autistic behaviors and epilepsy. *Clin Genet*. 2015;87(4):356–361.
76. Manole A, Mannikko R, Hanna MG, Kullmann DM, Houlden H. De novo KCNA2 mutations cause hereditary spastic paraplegia. *Ann Neurol*. 2017;81(2):326–328.
77. Saitsu H, Akita T, Tohyama J, et al. De novo KCNB1 mutations in infantile epilepsy inhibit repetitive neuronal firing. *Sci Rep*. 2015;5:15199.
78. Torkamani A, Bersell K, Jorge BS, et al. De novo KCNB1 mutations in epileptic encephalopathy. *Ann Neurol*. 2014;76(4):529–540.
79. Hoffjan S, Ibisler A, Tschentscher A, Dekomien G, Bidinost C, Rosa AL. WDR45 mutations in Rett (-like) syndrome and developmental delay: case report and an appraisal of the literature. *Mol Cell Probes*. 2016;30(1):44–49.
80. Lee JS, Yoo Y, Lim BC, et al. GM3 synthase deficiency due to ST3GAL5 variants in two Korean female siblings: masquerading as Rett syndrome-like phenotype. *Am J Med Genet A*. 2016;170(8):2200–2205.
81. Lambert L, Bienvenu T, Allou L, et al. MEF2C mutations are a rare cause of Rett or severe Rett-like encephalopathies. *Clin Genet*. 2012;82(5):499–501.
82. Craiu D, Dragostin O, Dica A, et al. Rett-like onset in late-infantile neuronal ceroid lipofuscinosis (CLN7) caused by compound heterozygous mutation in the MFSD8 gene and review of the literature data on clinical onset signs. *Eur J Paediatr Neurol*. 2015;19(1):78–86.
83. Ohba C, Nabatame S, Iijima Y, et al. De novo WDR45 mutation in a patient showing clinically Rett syndrome with childhood iron deposition in brain. *J Hum Genet*. 2014;59(5):292–295.
84. Gupta A. STXBP1-related EOEE- Early Onset Epilepsy AND Encephalopathy, or is it Early Onset Epileptic Encephalopathy?. *Epilepsy Curr*. 2016;16(5):302–304.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Schönewolf-Greulich B, Bisgaard A, Møller RS, et al. Clinician's guide to genes associated with Rett-like phenotypes—Investigation of a Danish cohort and review of the literature. *Clin Genet*. 2019;95:221–230. https://doi.org/10.1111/cge.13153

Book Chapter 1

Kaur S, Christodoulou J. MECP2 Disorders. 2001 Oct 3 [Updated 2019 Sep 19]. In: Adam M.P., Ardinger H.H., Pagon R.A., et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020.

U.S. National Library of Medicine
National Center for Biotechnology Information

NLM Citation: Kaur S, Christodoulou J. *MECP2* Disorders. 2001 Oct 3 [Updated 2019 Sep 19]. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020.
Bookshelf URL: https://www.ncbi.nlm.nih.gov/books/

MECP2 Disorders

Simranpreet Kaur, MSc, M Phil[1] and John Christodoulou, MBBS, PhD, FRACP, FFSc, FRCPA, FAHMS[2]

Created: October 3, 2001; Updated: September 19, 2019.

Summary

Clinical characteristics

The spectrum of *MECP2*-related phenotypes in females ranges from classic Rett syndrome to variant Rett syndrome with a broader clinical phenotype (either milder or more severe than classic Rett syndrome) to mild learning disabilities; the spectrum in males ranges from severe neonatal encephalopathy to pyramidal signs, parkinsonism, and macroorchidism (PPM-X) syndrome to severe syndromic/nonsyndromic intellectual disability.

- *Females*: Classic Rett syndrome, a progressive neurodevelopmental disorder primarily affecting girls, is characterized by apparently normal psychomotor development during the first six to 18 months of life, followed by a short period of developmental stagnation, then rapid regression in language and motor skills, followed by long-term stability. During the phase of rapid regression, repetitive, stereotypic hand movements replace purposeful hand use. Additional findings include fits of screaming and inconsolable crying, autistic features, panic-like attacks, bruxism, episodic apnea and/or hyperpnea, gait ataxia and apraxia, tremors, seizures, and acquired microcephaly.
- *Males:* Severe neonatal-onset encephalopathy, the most common phenotype in affected males, is characterized by a relentless clinical course that follows a metabolic-degenerative type of pattern, abnormal tone, involuntary movements, severe seizures, and breathing abnormalities. Death often occurs before age two years.

Diagnosis/testing

The diagnosis of a *MECP2* disorder is established by molecular genetic testing in a female proband with suggestive findings and a heterozygous *MECP2* pathogenic variant, and in a male proband with suggestive findings and a hemizygous *MECP2* pathogenic variant.

Author Affiliations: 1 Brain and Mitochondrial Research Group, Murdoch Children's Research Institute; Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia; Email: simran.kaur@mcri.edu.au. 2 Professor, Pediatrics and Biochemical, Molecular, and Human Genetics, Theme Director, Genetics Research, Murdoch Children's Research Institute; Chair of Genomic Medicine, Department of Pediatrics, University of Melbourne, Melbourne, Victoria, Australia; Email: john.christodoulou@mcri.edu.au.

Management

Treatment of manifestations: Treatment is mainly symptomatic and focuses on optimizing the individual's abilities using a multidisciplinary approach that should also include psychosocial support for family members. Risperidone may help in treating agitation; melatonin can ameliorate sleep disturbances. Treatment of seizures, constipation, gastroesophageal reflux, scoliosis, prolonged QTc, and spasticity as per standard care.

Surveillance: Periodic evaluation by the multidisciplinary team; regular assessment of QTc for evidence of prolongation; regular assessment for scoliosis.

Agents/circumstances to avoid: Drugs known to prolong the QT interval.

Genetic counseling

MECP2 disorders are inherited in an X-linked manner. More than 99% are simplex cases (i.e., a single occurrence in a family), resulting from a *de novo* pathogenic variant or possibly from inheritance of the pathogenic variant from a parent who has germline mosaicism. Rarely, a *MECP2* variant may be inherited from a heterozygous mother in whom favorable skewing of X-chromosome inactivation results in minimal to no clinical findings. When the mother is a known heterozygote, the risk to her offspring of inheriting the *MECP2* variant is 50%. When the pathogenic *MECP2* variant has been identified in the family, heterozygote testing for at-risk female relatives, prenatal testing for pregnancies at increased risk, and preimplantation genetic diagnosis are possible. Because of the possibility of parental germline mosaicism, it is appropriate to offer prenatal diagnosis to couples who have had a child with a *MECP2* disorder regardless of whether the *MECP2* pathogenic variant has been detected in a parent.

GeneReview Scope

MECP2 Disorders: Included Phenotypes [1, 2]	
Females	• *MECP2* classic Rett syndrome • Variant Rett syndrome • Mild learning disabilities
Males	• *MECP2*-related severe neonatal encephalopathy • Pyramidal signs, parkinsonism, and macroorchidism (PPM-X) syndrome • Syndromic/nonsyndromic intellectual disability

1. For other genetic causes of these phenotypes see Differential Diagnosis.
2. Note: The allelic disorder *MECP2* duplication syndrome is not included in this *GeneReview*. See Genetically Related Disorders.

Diagnosis

Note: Duplication of *MECP2* (ranging from 0.3 to 4 Mb and larger) is associated with the allelic disorder *MECP2* duplication syndrome and is not addressed in this *GeneReview*.

Suggestive Findings in Females

A *MECP2* disorder **should be suspected/considered** in females with the following clinical findings suggestive of *MECP2* classic Rett syndrome or variant Rett syndrome (based on clinical diagnostic criteria published by Neul et al [2010] [full text] prior to the widespread availability of molecular genetic testing), or mild learning disabilities.

Clinical findings of *MECP2* classic Rett syndrome and variant Rett syndrome

- Most distinguishing finding: A period of regression (range: ages 1-4 years) followed by recovery or stabilization (range: ages 2-10 years; mean: age 5 years)
- Main findings
  - Partial or complete loss of acquired purposeful hand skills
  - Partial or complete loss of acquired spoken language or language skill (e.g., babble)
  - Gait abnormalities: impaired (dyspraxic) or absence of ability
  - Stereotypic hand movements including hand wringing/squeezing, clapping/tapping, mouthing, and washing/rubbing automatisms
- Supportive findings
  - Breathing disturbances when awake
  - Bruxism when awake
  - Impaired sleep pattern
  - Abnormal muscle tone
  - Peripheral vasomotor disturbances
  - Scoliosis/kyphosis
  - Growth retardation
  - Small, cold hands and feet
  - Inappropriate laughing/screaming spells
  - Diminished response to pain
  - Intense eye communication - "eye pointing"
- Exclusionary findings
  - Brain injury secondary to peri- or postnatal trauma, neurometabolic disease, or severe infection that causes neurologic problems
  - Grossly abnormal psychomotor development in the first six months of life, with early milestones not being met

Clinical findings of *MECP2* mild learning disability. Typically mild and non-progressive. Note: Typically, females with mild learning disability are identified through molecular genetic testing following diagnosis of a first-degree relative (e.g., a more significantly affected brother or sister).

Suggestive Findings in Males

MECP2 disorders should be considered in a male with severe neonatal encephalopathy; pyramidal signs, parkinsonism, and macroorchidism (PPM-X) syndrome; or syndromic/nonsyndromic intellectual disability.

Clinical findings of *MECP2* severe neonatal encephalopathy

- Microcephaly
- Relentless clinical course that follows a metabolic-degenerative type of pattern
- Abnormal tone
- Involuntary movements
- Severe seizures
- Breathing abnormalities (including central hypoventilation or respiratory insufficiency)

Clinical findings of *MECP2* severe intellectual disability (including PPM-X syndrome)

- Moderate-to-severe intellectual disability
- Resting tremor
- Slowness of movements
- Ataxia
- PPM-X syndrome: *p*yramidal signs, *p*arkinsonism, and *m*acroorchidism

- No seizures or microcephaly
- Usually normal brain MRI, EEG, EMG, and nerve conduction velocity studies

Establishing the Diagnosis

Female proband. The diagnosis of a *MECP2* disorder **is usually established** in a female proband with suggestive findings and a heterozygous pathogenic variant in *MECP2* identified by molecular genetic testing (see Table 1).

Male proband. The diagnosis of a *MECP2* disorder **is established** in a male proband with suggestive findings and a hemizygous pathogenic variant in *MECP2* identified by molecular genetic testing (see Table 1).

Molecular genetic testing approaches can include a combination of **gene-targeted testing** (either single-gene or multigene panel) or **comprehensive genomic testing** (exome sequencing, exome array, genome sequencing) depending on the phenotype.

Gene-targeted testing requires that the clinician determine which gene(s) are likely involved, whereas genomic testing does not. Because the phenotype of *MECP2* disorders is broad, females with the distinctive findings described in Suggestive Findings are likely to be diagnosed using gene-targeted testing (see Option 1), whereas females and males with a phenotype indistinguishable from many other inherited disorders with intellectual disability and/or neonatal encephalopathy are more likely to be diagnosed using genomic testing (see Option 2).

Option 1

When the clinical findings suggest the diagnosis of a *MECP2* disorder, molecular genetic testing approaches can include use of single-gene testing or a **multigene panel**:

- **Single-gene testing.** Sequence analysis of *MECP2* detects small intragenic deletions/insertions and missense, nonsense, and splice site variants. If no pathogenic variant is found, perform gene-targeted deletion/duplication analysis to detect intragenic deletions or duplications. Note: Lack of amplification by PCR prior to sequence analysis can suggest a putative (multi)exon or whole-gene deletion on the X chromosome in affected males; confirmation requires additional testing by gene-targeted deletion/duplication analysis.
- Various **multigene panels** such as Rett/Angelman syndrome panels and more comprehensive childhood-onset epilepsy panels that include *MECP2* and other genes of interest (see Differential Diagnosis) are most likely to identify the genetic cause of the condition at the most reasonable cost while limiting identification of variants of uncertain significance and pathogenic variants in genes that do not explain the underlying phenotype. Note: (1) The genes included in the panel and the diagnostic sensitivity of the testing used for each gene vary by laboratory and are likely to change over time. (2) Some multigene panels may include genes not associated with the condition discussed in this *GeneReview*. (3) In some laboratories, panel options may include a custom laboratory-designed panel and/or custom phenotype-focused exome analysis that includes genes specified by the clinician. (4) Methods used in a panel may include sequence analysis, deletion/duplication analysis, and/or other non-sequencing-based tests. For this disorder a multigene panel that also includes deletion/duplication analysis is recommended (see Table 1).

  For an introduction to multigene panels click here. More detailed information for clinicians ordering genetic tests can be found here.

Option 2

When the phenotype overlaps with many other inherited disorders characterized by intellectual disability and/or neonatal encephalopathy, **comprehensive genomic testing** (which does not require the clinician to determine which gene[s] are likely involved) is another option. **Exome sequencing** is most commonly used; **genome sequencing** is also possible.

If exome sequencing is not diagnostic, **exome array** (when clinically available) may be considered to detect (multi)exon deletions or duplications that cannot be detected by sequence analysis.

For an introduction to comprehensive genomic testing click here. More detailed information for clinicians ordering genomic testing can be found here.

Table 1. Molecular Genetic Testing Used in *MECP2* Disorders

Gene [1]	Method	Proportion of Probands with a Pathogenic Variant [2] Detectable by Method
MECP2	Sequence analysis [3, 4]	90%-95% [5]
	Gene-targeted deletion/duplication analysis [6]	5%-10% [7, 8]

1. See Table A. Genes and Databases for chromosome locus and protein.
2. See Molecular Genetics for information on allelic variants detected in this gene.
3. Sequence analysis detects variants that are benign, likely benign, of uncertain significance, likely pathogenic, or pathogenic. Pathogenic variants may include small intragenic deletions/insertions and missense, nonsense, and splice site variants; typically, exon or whole-gene deletions/duplications are not detected. For issues to consider in interpretation of sequence analysis results, click here.
4. Lack of amplification by PCR prior to sequence analysis can suggest a putative (multi)exon or whole-gene deletion on the X chromosome in affected males; confirmation requires additional testing by gene-targeted deletion/duplication analysis.
5. Archer et al [2006], Philippe et al [2006]
6. Gene-targeted deletion/duplication analysis detects intragenic deletions or duplications. Methods used may include quantitative PCR, long-range PCR, multiplex ligation-dependent probe amplification (MLPA), and a gene-targeted microarray designed to detect single-exon deletions or duplications. Gene-targeted deletion/duplication testing will detect deletions ranging from a single exon to the whole gene; however, breakpoints of large deletions and/or deletion of adjacent genes (e.g., those described by Hardwick et al [2007]) may not be detected by these methods.
7. The sizes of many reported disease-associated deletions are at the upper limits of detection by sequence analysis and the lower limits of detection by gene-targeted deletion/duplication analysis; therefore, the proportion of pathogenic variants detected by either method depends on the methods used by a laboratory.
8. Archer et al [2006], Pan et al [2006], Philippe et al [2006], Hardwick et al [2007], Zahorakova et al [2007]

Clinical Characteristics

Clinical Description

In females the spectrum of *MECP2*-related phenotypes ranges from classic Rett syndrome, to variant Rett syndrome (either milder or more severe than classic Rett syndrome), to mild learning disabilities. In males the spectrum ranges from severe neonatal encephalopathy, to pyramidal signs, parkinsonism, and macroorchidism (PPM-X) syndrome, to severe syndromic/nonsyndromic intellectual disability.

MECP2 Disorders in Females

Table 2. Features of MECP2 Disorders in Females

Phenotype	Feature	% of Persons with Feature
***MECP2* classic Rett syndrome**	Regression followed by recovery or stabilization	99%
	Deceleration of head growth	80%
	Gait abnormalities	99%
	Seizures	60%-80%
	Hand stereotypies & loss of purposeful hand skills	100% [1]
	Absence of speech; high-pitched crying	99%
	Cold extremities	99%
	Irregular breathing	99%
Variant Rett syndrome	Regression followed by recovery or stabilization	99%
	Gait abnormalities	80%-99%
	Sleep disturbences	80%-99%
	seizures	6%-80%
	Hand stereotypies & loss of purposeful hand skills	97.3%
	Breathing irregularities	80%-99%
	Agitation	80%-99%

Gold et al [2018], Einspieler & Marschik [2019], Stallworth et al [2019]
1. Stallworth et al [2019]; 44% showed different patterns including hand wringing, washing, clapping, and tapping.

***MECP2* classic Rett syndrome.** Most individuals with classic Rett syndrome are female; however, males meeting the clinical criteria for classic Rett syndrome who have a 47,XXY karyotype [Hoffbuhr et al 2001, Leonard et al 2001, Schwartzman et al 2001] and postzygotic *MECP2* variants resulting in somatic mosaicism have been described [Clayton-Smith et al 2000, Topçu et al 2002].

Although early development is reportedly normal in children with classic Rett syndrome, parents – in retrospect – often identify subtle differences compared to unaffected sibs. Most (but not all) affected children have acquired microcephaly; stereotypic hand movements and breathing irregularities are seen in the majority.

Variant Rett syndrome. Females with variant Rett syndrome exhibit a broader spectrum of clinical features than those observed in classic Rett syndrome. At the more severe end of the spectrum, development is delayed from very early infancy; congenital hypotonia and infantile spasms are also seen. At the milder end of the spectrum, regression is less dramatic and intellectual disability is much less severe; some speech may be preserved.

Mild learning disabilities. In rare instances, females with a pathogenic *MECP2* variant may only exhibit mild learning disabilities or some autistic features, presumably as a consequence of favorable skewing of X-chromosome inactivation. When there is no regression phase and no characteristic hand stereotypes, the clinical course differs from that of classic and variant Rett syndrome.

MECP2 Disorders in Males

Table 3. Features of *MECP2* Disorders in Males

Phenotype	Feature	% of Persons with Feature		
		Present	Absent	Not reported
MECP2-related severe neonatal encephalopathy [1]	Normal birth parameters	71%		29%
	Head growth deceleration / microcephaly	94%		5.8%
	Hypotonia &/or feeding difficulties in infancy	82.4%		17.6%
	Hypertonia of extremities	52.9%	11.8%	35.3%
	Movement disorder, e.g., myoclonus, tremors, & dystonia	58.8%	17.7%	23.5%
	Mild cerebral atrophy	18%	35%	47%
	Polymicrogyria	5.9%	23.5%	70.6%
	Poor head control	35%	12%	53%
	Seizures	58.8%	17.7%	23.5%
	Severe development delay	82.4%		17.6%
	Irregular breathing / sleep apnea	47.1%	29.4%	23.5%
	Gastroesophageal reflux	35.3%		64.7%
	EEG abnormality	88.2%	5.9%	5.9%
Pyramidal signs, parkinsonism, and macroorchidism (PPM-X syndrome) [2]	Psychosis	67.6%	10.8%	21.6%
	Pyramidal signs	46%	2.7%	51.3%
	Macroorchidism	19%		81%
	Intellectual disability	50%		50%
	Parkinsonism	2.7%		97.3%
	Progressive spasticity	67.6%		32.4%
	Delayed development	54%		46%
	Speech difficulties	50%		50%
	Seizures	2.7%		
	Bilateral juvenile cataract	2.7%		
	Scoliosis or kyphosis	10.8%		
	Large ears	8.1%		
	Movement disorders	32.4%		
	Apraxia	2.7%		36%
	Seizures	8.1%		91.9%
	Dysmorphic features	5.4%		94.6%

Table 3. continued from previous page.

Phenotype	Feature	% of Persons with Feature		
		Present	Absent	Not reported
Syndromic/ nonsyndromic intellectual disability [3]	Severe intellectual disability	90%		10%
	Gait abnormalities	57%	7%	36%
	Facial dysmorphism	10%	3%	87%
	Behavioral problems	40%	3%	57%
	Autistic-like behavior	3%	53%	44%
	Seizures	20%	30%	50%
	Poor/absent language skills	47%	17%	36%
	Hypotonia	23%		77%
	Microcephaly	13%	23%	64%
	History of regression	17%	27%	56%
	Spasticity	33%	17%	50%
	Sleep disturbances	13%	10%	77%

1. Schüle et al [2008]
2. Lindsay et al [1996], Claes et al [1997], Gendrot et al [1999], Orrico et al [2000], Klauck et al [2002], Winnepenninckx et al [2002], Moog et al [2003], Psoni et al [2010]
3. Lubs et al [1999], Meloni et al [2000], Orrico et al [2000], Gomot et al [2003], Meins et al [2005], Van Esch et al [2005], Ramocki et al [2009]

Severe neonatal-onset encephalopathy is characterized by a relentless clinical course that follows a metabolic-degenerative type of pattern, abnormal tone, involuntary movements, severe seizures, and breathing abnormalities (including central hypoventilation or respiratory insufficiency) [Wan et al 1999, Villard et al 2000, Zeev et al 2002, Kankirawatana et al 2006]. Often, males with a severe neonatal encephalopathy die before age two years [Schanen et al 1998, Wan et al 1999].

The severe encephalopathy phenotype appears to be rare in females [Lugtenberg et al 2009].

X-linked ID and PPM-X syndrome. PPM-X syndrome, caused by the p.(Ala140Val) *MECP2* variant in males, is characterized by moderate-to-severe intellectual disability. Most have spasticity that may be progressive; some may have extrapyramidal movements. Episodic psychosis is seen in many but not all. Most affected males also have macroorchidism. Microcephaly is variable. See also Genotype-Phenotype Correlations.

Genotype-Phenotype Correlations

Genotype-phenotype correlations are inconsistent, due in part to the pattern of X-chromosome inactivation (XCI); females who have a *MECP2* pathogenic variant and favorably skewed XCI may have mild or no manifestations [Wan et al 1999, Amir et al 2000, Cheadle et al 2000, Huppke et al 2000, Weaving et al 2003, Chae et al 2004, Schanen et al 2004,Charman et al 2005].

MECP2 pathogenic variants with some residual function that are associated with milder phenotypes include the following:

- p.(Ala140Val). The phenotype is syndromic (PPM-X) intellectual disability in males and very mild cognitive impairment in females [Dotti et al 2002, Klauck et al 2002, Gomot et al 2003, Venkateswaran et al 2014, Lambert et al 2016, Sheikh et al 2016].

- p.(Arg133Cys). The phenotype is less severe than classic Rett syndrome in females; this variant can be present in affected males [Leonard et al 2003, Sheikh et al 2016].
- p.(Arg309Cys) is found in females and males with intellectual disability and some features of *MECP2* disorders, but not classic or variant Rett syndrome [Campos et al 2007, Schönewolf-Greulich et al 2016].

Prevalence

The worldwide prevalence is 1:10,000-1:23,000 female births [Ellaway et al 1999, Armstrong et al 2010]. Reports of incidence are limited; available estimates range from 0.43-0.71:10,000 for females in France [Bienvenu et al 2006] to 0.586:10,000 for females in Serbia [Sarajlija et al 2015] and 1.09:10,000 for females in Australia [Laurvick et al 2006].

Genetically Related (Allelic) Disorders

***MECP2* duplication syndrome** is characterized in affected males by infantile hypotonia, delayed psychomotor development leading to severe intellectual disability, poor speech development, progressive spasticity, recurrent respiratory infections, and seizures.

Duplications of *MECP2* ranging from 0.3 to 4 Mb and larger are found in all affected males.

The birth prevalence of *MECP2* duplication syndrome has been reported to be 0.65:100,000 for all live births and 1:100,000 for males in Australia with the median age at diagnosis of 23.5 months [Giudice-Nairn et al 2019].

Differential Diagnosis

Table 4. Disorders to Consider in the Differential Diagnosis of *MECP2* Disorders

Differential Diagnosis Disorder	Gene(s) / Genetic Mechanism	MOI	Clinical Features of Differential Diagnosis Disorder	
			Overlapping w/*MECP2* Disorders	Distinguishing from *MECP2* Disorders
Angelman syndrome	Deficient expression or function of maternally inherited *UBE3A* allele	See footnote 1	ID, severe speech impairment, gait ataxia &/or tremulousness of the limbs; microcephaly & seizures common; DD 1st noted at age ~6 mos	In classic Rett syndrome DD is not overtly evident in the 1st 6 mos.
Early infantile epileptic encephalopathy (OMIM 300672)	*CDKL5*	XL	In females: early-onset severe seizures w/poor cognitive development; facial gestalt, cortical visual impairment; In males: severe-profound ID & early-onset intractable seizures [2]	Very early-onset seizures, facial dysmorphism, & cortical visual impairment are not generally seen in classic Rett syndrome.

Table 4. continued from previous page.

Differential Diagnosis Disorder	Gene(s) / Genetic Mechanism	MOI	Clinical Features of Differential Diagnosis Disorder	
			Overlapping w/*MECP2* Disorders	Distinguishing from *MECP2* Disorders
Rett syndrome, congenital variant (OMIM 613454)	*FOXG1*	AD	Short normal period of development before onset of regression leading to severe ID, DD, postnatal microcephaly, agenesis of the corpus callosum, seizures, dyskinesia, & hypotonia [3]	Except for microcephaly, structural abnormalities are not usually seen on brain MRI.

AD = autosomal dominant; DD = developmental delay; ID = intellectual disability; MOI = mode of inheritance; XL = X-linked

1. The risk to sibs of a proband depends on the genetic mechanism leading to the loss of *UBE3A* function: typically less than 1% risk for probands with a deletion or uniparental disomy (UPD), and as high as 50% for probands with an imprinting defect or a pathogenic variant of *UBE3A*.

2. Elia et al [2008]

3. Overlapping features and a similar facial appearance between individuals with *FOXG1* pathogenic variants has led to the suggestion that these individuals should be regarded as having *FOXG1* syndrome rather than a variant of Rett syndrome [Kortüm et al 2011].

Management

Evaluations Following Initial Diagnosis

To establish the extent of disease and needs in an individual diagnosed with a *MECP2* disorder, the evaluations summarized in Table 5 (if not performed as part of the evaluation that led to the diagnosis) are recommended.

Table 5. Recommended Evaluations Following Initial Diagnosis in Individuals with a *MECP2* Disorder

System/Concern	Evaluation	Comment
Constitutional	Measurement of height, weight, & head circumference	
Neurologic	Neurologic evaluation	To incl brain MRI; consider EEG / video monitoring if seizures are a concern.
Development	Developmental assessment	• Motor, adaptive, cognitive, & speech/language evaluation • Evaluation for early intervention / special education
Psychiatric/ Behavioral	Neuropsychiatric evaluation	In individuals age >12 mos: screening for behavior problems incl sleep disturbances, ADHD, anxiety, &/or traits suggestive of ASD
Musculoskeletal	Orthopedics, physical medicine & rehabilitation, PT, OT evaluation	To incl assessment of: • Gross motor & fine motor skills • Scoliosis • Mobility & activities of daily living & need for adaptive devices • Need for PT (to improve gross motor skills) &/or OT (to improve fine motor skills)
Gastrointestinal/ Feeding	Gastroenterology / nutrition / feeding team evaluation	To incl: • Evaluation of aspiration risk & nutritional status • History of constipation & GERD Consider need for gastric tube placement.
Respiratory	Overnight sleep studies	• Analysis for abnormalities of breathing regularity • Noninvasive assessment of pulmonary gas exchange

Table 5. continued from previous page.

System/Concern	Evaluation	Comment
Sleep disorder	Breathing monitoring using portable polygraphic screening devices	To assess occurrence of apnea & hypopnea
Cardiovascular	Cardiac evaluation	To assess for prolonged QTc
Osteopenia	Bone densitometry	To assess for osteopenia
Eyes	Ophthalmologic evaluation	To assess for reduced vision, abnormal ocular movement, strabismus
Hearing	Audiology evaluation	Assess for hearing loss
ENT/Mouth		
Genitourinary		
Integument	History & examination	Reduced perfusion of hands & feet (possible autonomic abnormalities)
Miscellaneous/ Other	Consultation w/clinical geneticist &/or genetic counselor	To incl genetic counseling
	Family supports/resources	Assess: • Use of community or online resources, e.g., Parent to Parent • Need for social work involvement for parental support • Need for home nursing referral

ADHD = attention-deficit/hyperactivity disorder; ASD = autism spectrum disorder; GERD=gastroesophageal reflux disease; OT = occupational therapy; PT = physical therapy

Treatment of Manifestations

Treatment needs to be individualized following an assessment of the affected individual's clinical problems and needs.

Management is symptomatic and focuses on optimizing the individual's abilities using a multidisciplinary approach with input from a pediatric or adult specialist physician, dietician, occupational therapist, speech therapist, music therapist, dentist, and other medical subspecialists as needed.

Table 6. Treatment of Manifestations in Individuals with a *MECP2* Disorder

Manifestation/ Concern	Treatment	Considerations/Other
DD/ID	See Developmental Delay / Intellectual Disability Management Issues.	
Epilepsy	Standardized treatment w/AEDs by an experienced neurologist	• Many AEDs may be effective; no one AED has been demonstrated effective specifically for this disorder. • Education of parents/caregivers [1]
Psychiatric/ Behavioral	Risperidone (low dose) or selective serotonin uptake inhibitors have been somewhat successful in treating agitation.	
Musculoskeletal	Scoliosis	As per guidelines [2]
Poor weight gain / failure to thrive	Feeding therapy; gastrostomy tube placement may be required for persistent feeding issues	Low threshold for clinical feeding evaluation &/or radiographic swallowing study when showing clinical signs or symptoms of dysphagia; nutritional guidelines are available. [3]

Table 6. continued from previous page.

Manifestation/ Concern	Treatment	Considerations/Other
Spasticity	Orthopedics / physical medicine & rehabilitation / PT / OT incl stretching to help avoid contractures & falls	Consider need for positioning & mobility devices, disability parking placard.
Sleep disorder	Melatonin can ameliorate sleep disturbances.	Chloral hydrate, hydroxyzine, or diphenhydramine may be used w/melatonin.
Abnormal vision &/or strabismus	Standard treatment(s) as recommended by ophthalmologist	Community vision services through early intervention or school district
Central visual impairment	No specific treatment; early intervention to help stimulate visual development	
Hearing	Hearing aids may be helpful as per otolaryngologist	Community hearing services through early intervention or school district
Gastrointestinal	• Constipation: stool softeners, prokinetics, osmotic agents, or laxatives as needed • GERD: anti-reflux agents, smaller & thickened feedings, & positioning	
Cardiovascular	Treatment for prolonged QTc	Under care of pediatric cardiologist
Osteopenia	Baseline densitometry; optimization of physical activity & calcium & vitamin D levels	Guidelines for management of bone health are available. [4]
Family/ Community	• Ensure appropriate social work involvement to connect families w/local resources, respite, & support • Care coordination to manage multiple subspecialty appointments, equipment, medications, & supplies	• Ongoing assessment for need of palliative care involvement &/or home nursing • Consider involvement in adaptive sports or Special Olympics.

AED = antiepileptic drug; DD = developmental delay; GERD = gastroesophageal reflux disease; ID = intellectual disability
1. Education of parents/caregivers regarding common seizure presentations is appropriate. For information on non-medical interventions and coping strategies for children diagnosed with epilepsy, see Epilepsy & My Child Toolkit.
2. Downs et al [2009]
3. Leonard et al [2013]
4. Jefferson et al [2016]

Developmental Disability / Intellectual Disability Management Issues

The following information represents typical management recommendations for individuals with developmental delay / intellectual disability in the United States; standard recommendations may vary from country to country.

Ages 0-3 years. Referral to an early intervention program is recommended for access to occupational, physical, speech, and feeding therapy as well as infant mental health services, special educators, and sensory impairment specialists. In the US, early intervention is a federally funded program available in all states that provides in-home services to target individual therapy needs.

Ages 3-5 years. In the US, developmental preschool through the local public school district is recommended. Before placement, an evaluation is made to determine needed services and therapies and an individualized education plan (IEP) is developed for those who qualify based on established motor, language, social, or cognitive delay. The early intervention program typically assists with this transition. Developmental preschool is center based; for children too medically unstable to attend, home-based services are provided.

All ages. Consultation with a developmental pediatrician is recommended to ensure the involvement of appropriate community, state, and educational agencies and to support parents in maximizing quality of life. Some issues to consider:

- Individualized education plan (IEP) services:
  - An IEP provides specially designed instruction and related services to children who qualify.
  - IEP services will be reviewed annually to determine if any changes are needed.
  - As required by special education law, children should be in the least restrictive environment feasible at school and included in general education as much as possible and when appropriate.
  - Vision and hearing consultants should be a part of the child's IEP team to support access to academic material.
  - PT, OT, and speech services will be provided in the IEP to the extent that the need affects the child's access to academic material. Beyond that, private supportive therapies based on the affected individual's needs may be considered. Specific recommendations regarding type of therapy can be made by a developmental pediatrician.
  - As a child enters teen years, a transition plan should be discussed and incorporated in the IEP. For those receiving IEP services, the public school district is required to provide services until age 21.
- A 504 plan (Section 504: a US federal statute that prohibits discrimination based on disability) can be considered for those who require accommodations or modifications such as front-of-class seating, assistive technology devices, classroom scribes, extra time between classes, modified assignments, and enlarged text.
- Developmental Disabilities Administration (DDA) enrollment is recommended. DDA is a public agency that provides services and support to qualified individuals. Eligibility differs by state but is typically determined by diagnosis and/or associated cognitive/adaptive disabilities.
- Families with limited income and resources may also qualify for supplemental security income (SSI) for their child with a disability.

Motor Dysfunction

Gross motor dysfunction

- Physical therapy is recommended to maximize mobility and to reduce the risk for later-onset orthopedic complications (e.g., contractures, scoliosis, hip dislocation).
- Consider use of durable medical equipment and positioning devices as needed (e.g., wheelchairs, walkers, bath chairs, orthotics, adaptive strollers).
- For muscle tone abnormalities including hypertonia or dystonia, consider involving appropriate specialists to aid in management of baclofen, tizanidine, Botox®, anti-parkinsonian medications, or orthopedic procedures.

Fine motor dysfunction. Occupational therapy is recommended for difficulty with fine motor skills that affect adaptive function such as feeding, grooming, dressing, and writing.

Oral motor dysfunction should be assessed at each visit and clinical feeding evaluations and/or radiographic swallowing studies should be obtained for choking/gagging during feeds, poor weight gain, frequent respiratory illnesses, or feeding refusal that is not otherwise explained. Assuming that the child is safe to eat by mouth, feeding therapy (typically by an occupational or speech therapist) is recommended to improve coordination or sensory-related feeding issues. Feeds can be thickened or chilled for safety. When feeding dysfunction is severe, an NG-tube or G-tube may be necessary.

Communication issues. Consider evaluation for alternative means of communication (e.g., Augmentative and Alternative Communication [AAC]) for individuals who have expressive language difficulties. An AAC evaluation can be completed by a speech-language pathologist who has expertise in the area. The evaluation will

consider cognitive abilities and sensory impairments to determine the most appropriate form of communication. AAC devices can range from low-tech, such as picture exchange communication, to high-tech, such as voice-generating devices. Contrary to popular belief, AAC devices do not hinder verbal development of speech and in many cases can improve it.

Social/Behavioral Concerns

Children may qualify for and benefit from interventions used in treatment of autism spectrum disorder, including applied behavior analysis (ABA). ABA therapy is targeted to the individual child's behavioral, social, and adaptive strengths and weaknesses and typically performed one on one with a board-certified behavior analyst.

Consultation with a developmental pediatrician may be helpful in guiding parents through appropriate behavior management strategies or providing prescription medications, such as medication used to treat attention-deficit/hyperactivity disorder (ADHD), when necessary.

Concerns about serious aggressive or destructive behavior can be addressed by a pediatric psychiatrist.

Surveillance

Many of the clinical features in females with atypical Rett syndrome (Table 2) evolve with age and hence should be reassessed every six to 12 months.

Table 7. Recommended Surveillance for Individuals with a *MECP2* Disorder

System/Concern	Evaluation	Frequency
Feeding	Measurement of growth parameters	At each multidisciplinary clinic visit; minimum annually
	Evaluation of nutritional status & safety of oral intake	
Gastrointestinal	Monitor for constipation.	
Respiratory	Monitor for evidence of aspiration, respiratory insufficiency.	
Neurologic	Monitor those w/seizures as clinically indicated.	
	Assess for new manifestations, e.g., seizures, changes in tone, movement disorders.	
Development	Monitor developmental progress & educational needs.	
Speech & language	Monitor communication skills.	
Psychiatric/ Behavioral	Behavioral assessment for anxiety, attention, & aggressive or self-injurious behavior	
Musculoskeletal	Physical medicine, OT/PT assessment of mobility, self-help skills	
	Monitor scoliosis.	
Cardiovascular	Monitor for prolonged QTc.	
Respiratory	Apnea/hyperventilation	
Miscellaneous/ Other	Assess family need for social work support (e.g., palliative/respite care, home nursing; other local resources) & care coordination.	

Agents/Circumstances to Avoid

Because individuals with *MECP2* disorders are at increased risk for life-threatening arrhythmias associated with a prolonged QT interval, avoidance of drugs known to prolong the QT interval, including the following, is recommended:

- Prokinetic agents (e.g., cisapride)

- Antipsychotics (e.g., thioridazine), tricyclic antidepressants (e.g., imipramine)
- Antiarrhythmics (e.g., quinidine, sotolol, amiodarone)
- Anesthetic agents (e.g., thiopental, succinylcholine)
- Antibiotics (e.g., erythromycin, ketoconazole)

Click here (pdf) for a more extensive list of drugs to avoid.

Evaluation of Relatives at Risk

See Genetic Counseling for issues related to testing of at-risk relatives for genetic counseling purposes.

Therapies Under Investigation

A number of clinical trials are currently under way, including observational studies, studies focused on improvement of language and communication skills, and drug trials.

For details see www.rettsyndrome.org.

Search ClinicalTrials.gov in the US and EU Clinical Trials Register in Europe for access to information on clinical studies for a wide range of diseases and conditions.

Genetic Counseling

Genetic counseling is the process of providing individuals and families with information on the nature, inheritance, and implications of genetic disorders to help them make informed medical and personal decisions. The following section deals with genetic risk assessment and the use of family history and genetic testing to clarify genetic status for family members. This section is not meant to address all personal, cultural, or ethical issues that individuals may face or to substitute for consultation with a genetics professional. —ED.

Mode of Inheritance

MECP2 disorders are inherited in an X-linked manner.

Risk to Family Members

Parents of a proband

- Approximately 99.5% of affected individuals represent simplex cases (i.e., a single occurrence in the family).
- Female proband. *MECP2* molecular genetic testing is recommended for both parents.
- Male proband. *MECP2* molecular genetic testing is recommended for the mother. (Note: The father of an affected male will not have a *MECP2* disorder nor will he be hemizygous for the *MECP2* pathogenic variant; therefore, he does not require further evaluation/testing.)
- The mother of a proband who is found to be heterozygous for a *MECP2* variant may have favorably skewed X-chromosome inactivation that results in her being unaffected or mildly affected.
- If the *MECP2* pathogenic variant found in the proband cannot be detected in the leukocyte DNA of either parent, possible explanations include a *de novo* pathogenic variant in the proband or germline mosaicism in a parent. Maternal and paternal germline mosaicism have been reported [Amir et al 1999, Zeev et al 2002, Mari et al 2005].
  - Maternal germline mosaicism was reported in one of nine pregnancies [Mari et al 2005, Venâncio et al 2007].
  - Paternal germline mosaicism was reported in five fathers of affected daugthers from 21 families [Zhang et al 2019].

Sibs of a proband. The risk to sibs depends on the genetic status of the parents:

- If the mother of the proband has a *MECP2* pathogenic variant, the chance of transmitting it in each pregnancy is 50%.
  - Females who inherit the pathogenic variant are at high risk of developing a *MECP2* disorder, although skewed X-chromosome inactivation may result in a variable phenotype.
  - Males who inherit the variant may have a severe neonatal encephalopathy or, if they survive the first year, will most likely have a severe intellectual disability syndrome.
- If the proband represents a simplex case (i.e., a single occurrence in a family) and if the *MECP2* pathogenic variant cannot be detected in the leukocyte DNA of either parent, the risk to sibs is greater than that of the general population because of the possibility of parental germline mosaicism [Amir et al 1999, Zeev et al 2002, Mari et al 2005, Venâncio et al 2007, Zhang et al 2019].

Offspring of a proband

- Each child of a female proband with a *MECP2* disorder has a 50% chance of inheriting the *MECP2* pathogenic variant. Females with more severe *MECP2* disorders do not reproduce; mildly affected females have reproduced.
- Males with a *MECP2* disorder are not known to reproduce.

Other family members. The risk to other family members depends on the genetic status of the proband's mother: if the mother is affected or has a pathogenic *MECP2* variant, her family members may be at risk.

Related Genetic Counseling Issues

First-degree female relatives. Once the pathogenic *MECP2* variant has been identified in a proband, it is appropriate to offer testing to all first-degree female relatives regardless of their clinical status. Apparently unaffected sisters of a female proband with a *MECP2* disorder may be heterozygous for the *MECP2* variant present in their sister but have few to no manifestations because of skewed X-chromosome inactivation. Genetic counseling should address this possibility as clinically unaffected sisters may be at risk of transmitting the pathogenic *MECP2* variant to their children.

Family planning

- The optimal time for determination of genetic risk and discussion of the availability of prenatal testing is before pregnancy.
- It is appropriate to offer genetic counseling (including discussion of potential risks to offspring and reproductive options) to young adults who are mildly affected or are at risk of having a pathogenic *MECP2* variant.

DNA banking is the storage of DNA (typically extracted from white blood cells) for possible future use. Because it is likely that testing methodology and our understanding of genes, allelic variants, and diseases will improve in the future, consideration should be given to banking DNA of affected individuals.

Prenatal Testing and Preimplantation Genetic Diagnosis

Once the *MECP2* pathogenic variant has been identified in an affected family member, prenatal testing for a pregnancy at increased risk and preimplantation genetic diagnosis are possible. Males with a *MECP2* variant who survive infancy will most likely have severe intellectual disability. The phenotype in a female with a *MECP2* variant is difficult to predict and can range from apparently normal to severely affected.

Note: Because parental germline mosaicism for a *MECP2* pathogenic variant has been reported in multiple families, it is appropriate to offer prenatal testing to the parents of a child with a *MECP2* disorder whether or not the *MECP2* pathogenic variant has been identified in the leukocyte DNA of either parent.

Resources

GeneReviews staff has selected the following disease-specific and/or umbrella support organizations and/or registries for the benefit of individuals with this disorder and their families. GeneReviews is not responsible for the information provided by other organizations. For information on selection criteria, click here.

- **International Rett Syndrome Foundation (IRSF)**
  4600 Devitt Drive
  Cincinnati OH 45246
  Phone: 800-818-7388 (toll-free); 513-874-3020
  Fax: 513-874-2520
  www.rettsyndrome.org
- **My46 Trait Profile**
  Rett syndrome
- **National Library of Medicine Genetics Home Reference**
  Rett syndrome
- **NCBI Genes and Disease**
  Rett syndrome
- **Rett New Zealand**
  PO Box 28 049
  Wellington
  New Zealand
  Phone: 04 475 9265
  Email: rett.info@nzord.org.nz
  www.rettsyndrome.org.nz
- **Rett Syndrome Europe**
  www.rettsyndrome.eu
- **Reverse Rett**
  United Kingdom
  www.reverserett.org.uk
- **Medical Home Portal**
  The Parents & Families section of the Medical Home Portal provides information and resources to help families learn how to better care for a child with chronic and complex conditions and to become more effective partners in their child's care.
  Department of Pediatrics University of Utah

P.O. Box 581289
Salt Lake City UT 84158
Phone: 801-213-3920
Email: info@medicalhomeportal.org
For Parents & Families

- **Rett Syndrome Research Trust**
  67 Under Cliff Road
  Trumbull CT 06611
  Phone: 203-445-0041
  Fax: 203-445-9234
  Email: monica@rsrt.org
  www.rsrt.org
- **Rett UK**
  Langham House West
  Mill Street
  Luton LU1 2NA
  United Kingdom
  Phone: 01582 798 910
  Email: info@rettuk.org
  www.rettuk.org
- **Australian Rett Syndrome Study / InterRett Registry**
  Telethon Institute for Child Health Research
  PO Box 855
  West Perth 6872
  Australia
  Phone: +61 8 9489 7790; +61 419 956 946
  Fax: +61 8 9489 7700
  Email: rett@ichr.uwa.edu.au
  www.aussierett.org.au/about-us.aspx

Molecular Genetics

Information in the Molecular Genetics and OMIM tables may differ from that elsewhere in the GeneReview: tables may contain more recent information. —ED.

Table A. MECP2 Disorders: Genes and Databases

Gene	Chromosome Locus	Protein	Locus-Specific Databases	HGMD	ClinVar

Table A. continued from previous page.

MECP2	Xq28	Methyl-CpG-binding protein 2	MECP2 @ LOVD CCHMC - Human Genetics Mutation Database (MECP2) RettBASE	MECP2	MECP2

Data are compiled from the following standard references: gene from HGNC; chromosome locus from OMIM; protein from UniProt. For a description of databases (Locus Specific, HGMD, ClinVar) to which links are provided, click here.

Table B. OMIM Entries for MECP2 Disorders (View All in OMIM)

300005	METHYL-CpG-BINDING PROTEIN 2; MECP2
300055	MENTAL RETARDATION, X-LINKED, SYNDROMIC 13; MRXS13
300496	AUTISM, SUSCEPTIBILITY TO, X-LINKED 3; AUTSX3
300673	ENCEPHALOPATHY, NEONATAL SEVERE, DUE TO MECP2 MUTATIONS
312750	RETT SYNDROME; RTT

Molecular Pathogenesis

Loss of the protein MeCP2 leads to epigenetic aberrations of chromatin, suggesting that MeCP2 deficiency could lead to loss of imprinting, thereby contributing to the pathogenesis of Rett syndrome [Horike et al 2005, Kaufmann et al 2005, Makedonski et al 2005].

The nuclear MeCP2 protein functional domains include:

- Methyl binding domain (MBD): binds specifically to DNA at symmetrically methylated CpGs within chromatin [Hansen et al 2010, Casas-Delucchi et al 2012]
- Transcription repression domain (TRD): responsible for recruiting other proteins that mediate transcription repression
- A-T hook domain: basic residues that bind A-T rich DNA [Baker et al 2013, Heckman et al 2014]
- WW domain: conserved C-terminal domain [Buschdorf & Stratling 2004]

It has also been shown that MeCP2 plays a role in gene splicing [Young et al 2005] and in long-range chromatin remodeling [Horike et al 2005], and may be a transcriptional activator [Chahrour et al 2008].

Mechanism of disease causation. Most pathogenic *MECP2* variants occur *de novo*. It has been suggested that pathogenic variants result in loss of protein function; some functional studies show that pathogenic *MECP2* variants affect the MBD or TRD domains of the abnormal protein, depending on the location of the variant [Kudo et al 2001, Kudo et al 2002, Kudo et al 2003].

***MECP2*-specific laboratory technical considerations.** Two transcripts have been described:

- NM_001110792.1: encodes *MECP2_e1*, includes exons 1, 3, and 4 but not exon 2 (498 amino acids)
- NM_004992.3, encodes *MECP2_e2,* includes exons 2, 3, and 4 but not exon 1 (486 amino acids)

Although the isoforms are nearly identical, use of two alternative start codons creates alternative N-termini. The e1 transcript is much more highly expressed in brain than the e2 transcript [Kriaucionis & Bird 2004, Mnatzakanian et al 2004].Of note:

- Exon 1 (*MECP2_e1)*: pathogenic variants in exon 1 are rare and include variants in the start codon (p.Met1?) and p.Ala2 as well as variant frameshift changes [Amir et al 2005, Evans et al 2005, Poirier et al 2005, Ravn et al 2005, Saxena et al 2006, Saunders et al 2009, Sheikh et al 2017].

- Exon 2 (*MECP2_e2*): a pathogenic variant in the start codon (p.Met1?) has been reported in exon 2 [Gauthier et al 2005].

The majority of pathogenic variants occur in the region encoding the methyl binding domain (MBD, exons 3 and 4; amino acids 90-174 of the MeCP2 e2 isoform), affecting the ability of the MeCP2 protein to bind to target DNA. A number of highly recurrent nonsense variants are found in the transcriptional repression domain (TRD, exon 4; amino acids 219-322 of the MeCP2 e2 isoform) and beyond the TRD, a large number of frameshift variants delete the C-terminal end of the protein (3' end of exon 4).

Table 8. Notable *MECP2* Pathogenic Variants

Reference Sequences	DNA Nucleotide Change	Predicted Protein Change	Comment [Reference]
NM_004492.3 NP_004983.1	c.473C>T	p.(Thr158Met)	Common, recurrent pathogenic variants [Miltenberger-Miltenyi & Laccone 2003, Archer et al 2006, Philippe et al 2006]
	c.502C>T	p.(Arg168Ter)	
	c.763C>T	p.(Arg255Ter)	
	c.808C>T	p.(Arg270Ter)	
	c.916C>T	p.(Arg306Cys)	
	c.397C>T	p.(Arg133Cys)	Milder phenotype in females is consistent w/in vitro functional studies showing that DNA binding is not impaired [Leonard et al 2003, Sheikh et al 2016].
	c.419C>T	p.(Ala140Val)	Non-classic, variant Rett syndrome, observed in familial cases w/affected males [Dotti et al 2002, Klauck et al 2002, Gomot et al 2003, Venkateswaran et al 2014, Lambert et al 2016, Sheikh et al 2016]; heterozygous females may have mild ID & impaired speech acquistion [Klauck et al 2002, Lambert et al 2016].
	c.925C>T	p.(Arg309Trp)	Observed in females & males w/ID & some features of a *MECP2* disorder, but not classic or variant Rett syndrome [Campos et al 2007, Schönewolf-Greulich et al 2016]

ID = intellectual disability
Variants listed in the table have been provided by the authors. *GeneReviews* staff have not independently verified the classification of variants.
GeneReviews follows the standard naming conventions of the Human Genome Variation Society (varnomen.hgvs.org). See Quick Reference for an explanation of nomenclature.

References

Published Guidelines / Consensus Statements

Downs J, Bergman A, Carter P, Anderson A, Palmer GM, Roye D, van Bosse H, Bebbington A, Larsson EL, Smith BG, Baikie G, Fyfe S, Leonard H. Guidelines for management of scoliosis in Rett syndrome patients based on expert consensus and clinical evidence. Spine. 2009;34:E607–17. PubMed PMID: 19644320.

Literature Cited

Amir RE, Fang P, Yu Z, Glaze DG, Percy AK, Zoghbi HY, Roa BB, Van den Veyver IB. Mutations in exon 1 of MECP2 are a rare cause of Rett syndrome. J Med Genet. 2005;42:e15. PubMed PMID: 15689438.

Amir RE, Van den Veyver IB, Schultz R, Malicki DM, Tran CQ, Dahle EJ, Philippi A, Timar L, Percy AK, Motil KJ, Lichtarge O, Smith EO, Glaze DG, Zoghbi HY. Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. Ann Neurol. 2000;47:670–9. PubMed PMID: 10805343.

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl- CpG-binding protein 2. Nat Genet. 1999;23:185–8. PubMed PMID: 10508514.

Archer HL, Whatley SD, Evans JC, Ravine D, Huppke P, Kerr A, Bunyan D, Kerr B, Sweeney E, Davies SJ, Reardon W, Horn J, MacDermot KD, Smith RA, Magee A, Donaldson A, Crow Y, Hermon G, Miedzybrodzka Z, Cooper DN, Lazarou L, Butler R, Sampson J, Pilz DT, Laccone F, Clarke AJ. Gross rearrangements of the MECP2 gene are found in both classical and atypical Rett syndrome patients. J Med Genet. 2006;43:451–6. PubMed PMID: 16183801.

Armstrong AH, Hangauer J, Agazzi H, Nunez A, Gieron-Korthals M. Individuals with intellectual and developmental disabilities. In: David AS, ed. *Handbook of Pediatric Neuropsychology*. New York: Springer. 2010;537-50.

Baker SA, Chen L, Wilkins AD, Yu P, Lichtarge O, Zoghbi HY. An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. Cell. 2013;152:984–96. PubMed PMID: 23452848.

Bienvenu T, Philippe C, de Roux N, Raynaud M, Bonnefond JP, Pasquier L, Lesca G, Mancini J, Jonveaux P, Moncla A, Feingold J, Chelly J, Villard L. The incidence of Rett syndrome in France. Pediatr Neurol. 2006;34:372–5. PubMed PMID: 16647997.

Buschdorf JP, Stratling WH. A WW domain binding region in methyl-CpG-binding protein MeCP2: impact on Rett syndrome. J Mol Med. 2004;82:135–43. PubMed PMID: 14618241.

Campos M Jr, Abdalla CB, Santos-Rebouças CB, dos Santos AV, Pestana CP, Domingues ML, dos Santos JM, Pimentel MM. Low significance of MECP2 mutations as a cause of mental retardation in Brazilian males. Brain Dev. 2007;29:293–7. PubMed PMID: 17084570.

Casas-Delucchi CS, Becker A, Bolius JJ, Cristina Cardoso M. Targeted manipulation of heterochromatin rescues MeCP2 Rett mutants and re-establishes higher order chromatin organization. Nucl Acids Res. 2012;40:e176. PubMed PMID: 22923521.

Chae JH, Hwang H, Hwang YS, Cheong HJ, Kim KJ. Influence of MECP2 gene mutation and X-chromosome inactivation on the Rett syndrome phenotype. J Child Neurol. 2004;19:503–8. PubMed PMID: 15526954.

Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY. MeCP2, a key contributor to neurological disease, activates and represses transcription. Science. 2008;320:1224–9. PubMed PMID: 18511691.

Charman T, Neilson TC, Mash V, Archer H, Gardiner MT, Knudsen GP, McDonnell A, Perry J, Whatley SD, Bunyan DJ, Ravn K, Mount RH, Hastings RP, Hulten M, Orstavik KH, Reilly S, Cass H, Clarke A, Kerr AM, Bailey ME. Dimensional phenotypic analysis and functional categorisation of mutations reveal novel genotype-phenotype associations in Rett syndrome. Eur J Hum Genet. 2005;13:1121–30. PubMed PMID: 16077736.

Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A. Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location. Hum Mol Genet. 2000;9:1119–29. PubMed PMID: 10767337.

Claes S, Devriendt K, D'Adamo P, Meireleire J, Raeymaekers P, Toniolo D, Cassiman JJ, Fryns JP. X-linked severe mental retardation and a progressive neurological disorder in a Belgian family: clinical and genetic studies. Clin Genet. 1997;52:155–61. PubMed PMID: 9377804.

Clayton-Smith J, Watson P, Ramsden S, Black GC. Somatic mutation in MECP2 as a non-fatal neurodevelopmental disorder in males. Lancet. 2000;356:830–2. PubMed PMID: 11022934.

Dotti MT, Orrico A, De Stefano N, Battisti C, Sicurelli F, Severi S, Lam CW, Galli L, Sorrentino V, Federico A. A Rett syndrome MECP2 mutation that causes mental retardation in men. Neurology. 2002;58:226–30. PubMed PMID: 11805248.

Downs J, Bergman A, Carter P, Anderson A, Palmer GM, Roye D, van Bosse H, Bebbington A, Larsson EL, Smith BG, Baikie G, Fyfe S, Leonard H. Guidelines for management of scoliosis in Rett syndrome patients based on expert consensus and clinical evidence. Spine. 2009;34:E607–17. PubMed PMID: 19644320.

Einspieler C, Marschik PB. Regression in Rett syndrome: developmental pathways to its onset. Neurosci Biobehav Rev. 2019;98:320–32. PubMed PMID: 30832924.

Elia M, Falco M, Ferri R, Spalletta A, Bottitta M, Calabrese G, Carotenuto M, Musumeci SA, Lo Giudice M, Fichera M. CDKL5 mutations in boys with severe encephalopathy and early-onset intractable epilepsy. Neurology. 2008;71:997–9. PubMed PMID: 18809835.

Ellaway C, Williams K, Leonard H, Higgins G, Wilcken B, Christodoulou J. Rett syndrome: randomized controlled trial of L-carnitine. J Child Neurol. 1999;14:162–7. PubMed PMID: 10190267.

Evans JC, Archer HL, Whatley SD, Kerr A, Clarke A, Butler R. Variation in exon 1 coding region and promoter of MECP2 in Rett syndrome and controls. Eur J Hum Genet. 2005;13:124–6. PubMed PMID: 15367913.

Gauthier J, de Amorim G, Mnatzakanian GN, Saunders C, Vincent JB, Toupin S, Kauffman D, St-Onge J, Laurent S, Macleod PM, Minassian BA, Rouleau GA. Clinical stringency greatly improves mutation detection in Rett syndrome. Can J Neurol Sci. 2005;32:321–6. PubMed PMID: 16225173.

Gendrot C, Ronce N, Raynaud M, Ayrault AD, Dourlens J, Castelnau P, Muh JP, Chelly J, Moraine C. X-linked nonspecific mental retardation (MRX16) mapping to distal Xq2 8: linkage study and neuropsychological data in a large family. Am J Med Genet. 1999;83:411–8. PubMed PMID: 10232754.

Giudice-Nairn P, Downs J, Wong K, Wilson D, Ta D, Gattas M, Amor D, Thompson E, Kirrali-Borri C, Ellaway C, Leonard H. The incidence, prevalence and clinical features of MECP2 duplication syndrome in Australian children. J Paediatr Child Health. 2019. Epub ahead of print. PubMed PMID: 30756435.

Gold WA, Krishnaraj R, Ellaway C, Christodoulou J. Rett syndrome: a genetic update and clinical review focusing on comorbidities. ACS Chem Neurosci. 2018;9:167–76. PubMed PMID: 29185709.

Gomot M, Gendrot C, Verloes A, Raynaud M, David A, Yntema HG, Dessay S, Kalscheuer V, Frints S, Couvert P, Briault S, Blesson S, Toutain A, Chelly J, Desportes V, Moraine C. MECP2 gene mutations in non-syndromic X-linked mental retardation: phenotype-genotype correlation. Am J Med Genet A. 2003;123A:129–39. PubMed PMID: 14598336.

Hansen JC, Ghosh RP, Woodcock CL. Binding of the Rett syndrome protein, MeCP2, to methylated and unmethylated DNA and chromatin. IUBMB Life. 2010;62:732–8. PubMed PMID: 21031501.

Hardwick SA, Reuter K, Williamson SL, Vasudevan V, Donald J, Slater K, Bennetts B, Bebbington A, Leonard H, Williams SR, Smith RL, Cloosterman D, Christodoulou J. Delineation of large deletions of the MECP2 gene in Rett syndrome patients, including a familial case with a male proband. Eur J Hum Genet. 2007;15:1218–29. PubMed PMID: 17712354.

Heckman LD, Chahrour MH, Zoghbi HY. Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. eLife. 2014;3:e02676. PubMed PMID: 24970834.

Hoffbuhr K, Devaney JM, LaFleur B, Sirianni N, Scacheri C, Giron J, Schuette J, Innis J, Marino M, Philippart M, Narayanan V, Umansky R, Kronn D, Hoffman EP, Naidu S. MeCP2 mutations in children with and without the phenotype of Rett syndrome. Neurology. 2001;56:1486–95. PubMed PMID: 11402105.

Horike S, Cai S, Miyano M, Cheng JF, Kohwi-Shigematsu T. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet. 2005;37:31–40. PubMed PMID: 15608638.

Huppke P, Laccone F, Kramer N, Engel W, Hanefeld F. Rett syndrome: analysis of MECP2 and clinical characterization of 31 patients. Hum Mol Genet. 2000;9:1369–75. PubMed PMID: 10814718.

Jefferson A, Leonard H, Siafarikas A, Woodhead H, Fyfe S, Ward LM, Munns C, Motil K, Tarquinio D, Shapiro JR, Brismar T, Ben-Zeev B, Bisgaard AM, Coppola G, Ellaway C, Freilinger M, Geerts S, Humphreys P, Jones

M, Lane J, Larsson G, Lotan M, Percy A, Pineda M, Skinner S, Syhler B, Thompson S, Weiss B, Witt Engerström I, Downs J. Clinical guidelines for management of bone health in Rett syndrome based on expert consensus and available evidence. PLoS One. 2016;11:e0146824. PubMed PMID: 26849438.

Kankirawatana P, Leonard H, Ellaway C, Scurlock J, Mansour A, Makris CM, Dure LS 4th, Friez M, Lane J, Kiraly-Borri C, Fabian V, Davis M, Jackson J, Christodoulou J, Kaufmann WE, Ravine D, Percy AK. Early progressive encephalopathy in boys and MECP2 mutations. Neurology. 2006;67:164–6. PubMed PMID: 16832102.

Kaufmann WE, Jarrar MH, Wang JS, Lee YJ, Reddy S, Bibat G, Naidu S. Histone modifications in Rett syndrome lymphocytes: a preliminary evaluation. Brain Dev. 2005;27:331–9. PubMed PMID: 16023547.

Klauck SM, Lindsay S, Beyer KS, Splitt M, Burn J, Poustka A. A mutation hot spot for nonspecific X-linked mental retardation in the MECP2 gene causes the PPM-X syndrome. Am J Hum Genet. 2002;70:1034–7. PubMed PMID: 11885030.

Kortüm F, Das S, Flindt M, Morris-Rosendahl DJ, Stefanova I, Goldstein A, Horn D, Klopocki E, Kluger G, Martin P, Rauch A, Roumer A, Saitta S, Walsh LE, Wieczorek D, Uyanik G, Kutsche K, Dobyns WB. The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. J Med Genet. 2011;48:396–406. PubMed PMID: 21441262.

Kriaucionis S, Bird A. The major form of MeCP2 has a novel N-terminus generated by alternative splicing. Nucleic Acids Res. 2004;32:1818–23. PubMed PMID: 15034150.

Kudo S, Nomura Y, Segawa M, Fujita N, Nakao M, Dragich J, Schanen C, Tamura M. Functional analyses of MeCP2 mutations associated with Rett syndrome using transient expression systems. Brain Dev. 2001;23:S165–73. PubMed PMID: 11738866.

Kudo S, Nomura Y, Segawa M, Fujita N, Nakao M, Hammer S, Schanen C, Terai I, Tamura M. Functional characterisation of MeCP2 mutations found in male patients with X linked mental retardation. J Med Genet. 2002;39:132–6. PubMed PMID: 11836365.

Kudo S, Nomura Y, Segawa M, Fujita N, Nakao M, Schanen C, Tamura M. Heterogeneity in residual function of MeCP2 carrying missense mutations in the methyl CpG binding domain. J Med Genet. 2003;40:487–93. PubMed PMID: 12843318.

Lambert S, Maystadt I, Boulanger S, Vrielynck P, Destrée A, Lederer D, Moortgat S. Expanding phenotype of p.Ala140Val mutation in MECP2 in a 4 generation family with X-linked intellectual disability and spasticity. Eur J Med Genet. 2016;59:522–5. PubMed PMID: 27465203.

Laurvick CL, de Klerk N, Bower C, Christodoulou J, Ravine D, Ellaway C, Williamson S, Leonard H. Rett syndrome in Australia: a review of the epidemiology. J Pediatr. 2006;148:347–52. PubMed PMID: 16615965.

Leonard H, Colvin L, Christodoulou J, Schiavello T, Williamson S, Davis M, Ravine D, Fyfe S, de Klerk N, Matsuishi T, Kondo I, Clarke A, Hackwell S, Yamashita Y. Patients with the R133C mutation: is their phenotype different from patients with Rett syndrome with other mutations? J Med Genet. 2003;40:e52. PubMed PMID: 12746406.

Leonard H, Ravikumara M, Baikie G, Naseem N, Ellaway C, Percy A, Abraham S, Geerts S, Lane J, Jones M, Bathgate K, Downs J, et al. Assessment and management of nutrition and growth in Rett syndrome. J Pediatr Gastroenterol Nutr. 2013;57:451–60. PubMed PMID: 24084372.

Leonard H, Silberstein J, Falk R, Houwink-Manville I, Ellaway C, Raffaele LS, Engerstrom IW, Schanen C. Occurrence of Rett syndrome in boys. J Child Neurol. 2001;16:333–8. PubMed PMID: 11392517.

Lindsay S, Splitt M, Edney S, Berney TP, Knight SJ, Davies KE, O'Brien O, Gale M, Burn J. PPM-X: a new X-linked mental retardation syndrome with psychosis, pyramidal signs, and macroorchidism maps to Xq28. Am J Hum Genet. 1996;58:1120–6. PubMed PMID: 8651288.

Lubs H, Abidi F, Bier JA, Abuelo D, Ouzts L, Voeller K, Fennell E, Stevenson RE, Schwartz CE, Arena F. XLMR syndrome characterized by multiple respiratory infections, hypertelorism, severe CNS deterioration and early death localizes to distal Xq28. Am J Med Genet. 1999;85:243–8. PubMed PMID: 10398236.

Lugtenberg D, Kleefstra T, Oudakker AR, Nillesen WM, Yntema HG, Tzschach A, Raynaud M, Rating D, Journel H, Chelly J, Goizet C, Lacombe D, Pedespan JM, Echenne B, Tariverdian G, O'Rourke D, King MD, Green A, van Kogelenberg M, Van Esch H, Gecz J, Hamel BC, van Bokhoven H, de Brouwer AP. Structural variation in Xq28: MECP2 duplications in 1% of patients with unexplained XLMR and in 2% of male patients with severe encephalopathy. Eur J Hum Genet. 2009;17:444–53. PubMed PMID: 18985075.

Makedonski K, Abuhatzira L, Kaufman Y, Razin A, Shemer R. MeCP2 deficiency in Rett syndrome causes epigenetic aberrations at the PWS/AS imprinting center that affects UBE3A expression. Hum Mol Genet. 2005;14:1049–58. PubMed PMID: 15757975.

Mari F, Caselli R, Russo S, Cogliati F, Ariani F, Longo I, Bruttini M, Meloni I, Pescucci C, Schurfeld K, Toti P, Tassini M, Larizza L, Hayek G, Zappella M, Renieri A. Germline mosaicism in Rett syndrome identified by prenatal diagnosis. Clin Genet. 2005;67:258–60. PubMed PMID: 15691364.

Meins M, Lehmann J, Gerresheim F, Herchenbach J, Hagedorn M, Hameister K, Epplen JT. Submicroscopic duplication in Xq28 causes increased expression of the MECP2 gene in a boy with severe mental retardation and features of Rett syndrome. J Med Genet. 2005;42:e12. PubMed PMID: 15689435.

Meloni I, Bruttini M, Longo I, Mari F, Rizzolio F, D'Adamo P, Denvriendt K, Fryns JP, Toniolo D, Renieri A. A mutation in the rett syndrome gene, MECP2, causes X-linked mental retardation and progressive spasticity in males. Am J Hum Genet. 2000;67:982–5. PubMed PMID: 10986043.

Miltenberger-Miltenyi G, Laccone F. Mutations and polymorphisms in the human methyl CpG-binding protein MECP2. Hum Mutat. 2003;22:107–15. PubMed PMID: 12872250.

Mnatzakanian GN, Lohi H, Munteanu I, Alfred SE, Yamada T, MacLeod PJ, Jones JR, Scherer SW, Schanen NC, Friez MJ, Vincent JB, Minassian BA. A previously unidentified MECP2 open reading frame defines a new protein isoform relevant to Rett syndrome. Nat Genet. 2004;36:339–41. PubMed PMID: 15034579.

Moog U, Smeets EE, van Roozendaal KE, Schoenmakers S, Herbergs J, Schoonbrood-Lenssen AM, Schrander-Stumpel CT. Neurodevelopmental disorders in males related to the gene causing Rett syndrome in females (MECP2). Eur J Paediatr Neurol. 2003;7:5–12. PubMed PMID: 12615169.

Neul JL, Kaufmann WE, Glaze DG, Christodoulou J, Clarke AJ, Bahi-Buisson N, Leonard H, Bailey ME, Schanen NC, Zappella M, Renieri A, Huppke P, Percy AK, et al. Rett syndrome: revised diagnostic criteria and nomenclature. Ann Neurol. 2010;68:944–50. PubMed PMID: 21154482.

Orrico A, Lam C, Galli L, Dotti MT, Hayek G, Tong SF, Poon PM, Zappella M, Federico A, Sorrentino V. MECP2 mutation in male patients with non-specific X-linked mental retardation. FEBS Lett. 2000;481:285–8. PubMed PMID: 11007980.

Pan H, Li MR, Nelson P, Bao XH, Wu XR, Yu S. Large deletions of the MECP2 gene in Chinese patients with classical Rett syndrome. Clin Genet. 2006;70:418–9. PubMed PMID: 17026625.

Philippe C, Villard L, De Roux N, Raynaud M, Bonnefond JP, Pasquier L, Lesca G, Mancini J, Jonveaux P, Moncla A, Chelly J, Bienvenu T. Spectrum and distribution of MECP2 mutations in 424 Rett syndrome patients: a molecular update. Eur J Med Genet. 2006;49:9–18. PubMed PMID: 16473305.

Poirier K, Francis F, Hamel B, Moraine C, Fryns JP, Ropers HH, Chelly J, Bienvenu T. Mutations in exon 1 of MECP2B are not a common cause of X-linked mental retardation in males. Eur J Hum Genet. 2005;13:523–4. PubMed PMID: 15770224.

Psoni S, Sofocleous C, Traeger-Synodinos J, Kitsiou-Tzeli S, Kanavakis E, Fryssira-Kanioura H. Phenotypic and genotypic variability in four males with MECP2 gene sequence aberrations including a novel deletion. Pediatr Res. 2010;67:551–6. PubMed PMID: 20098342.

Ramocki MB, Peters SU, Tavyev YJ, Zhang F, Carvalho CM, Schaaf CP, Richman R, Fang P, Glaze DG, Lupski JR, Zoghbi HY. Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. Ann Neurol. 2009;66:771–82. PubMed PMID: 20035514.

Ravn K, Nielsen JB, Skjeldal OH, Kerr A, Hulten M, Schwartz M. Large genomic rearrangements in MECP2. Hum Mutat. 2005;25:324. PubMed PMID: 15712379.

Sarajlija A, Kisic-Tepavcevic D, Nikolic Z, Savic Pavicevic D, Obradovic S, Djuric M, Pekmezovic T. Epidemiology of Rett syndrome in Serbia: prevalence, incidence and survival. Neuroepidemiology. 2015;44:1–5. PubMed PMID: 25571926.

Saunders CJ, Minassian BE, Chow EW, Zhao W, Vincent JB. Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. Am J Med Genet A. 2009;149A:1019–23. PubMed PMID: 19365833.

Saxena A, de Lagarde D, Leonard H, Williamson SL, Vasudevan V, Christodoulou J, Thompson E, MacLeod P, Ravine D. Lost in translation: translational interference from a recurrent mutation in exon 1 of MECP2. J Med Genet. 2006;43:470–7. PubMed PMID: 16155192.

Schanen C, Houwink EJ, Dorrani N, Lane J, Everett R, Feng A, Cantor RM, Percy A. Phenotypic manifestations of MECP2 mutations in classical and atypical Rett syndrome. Am J Med Genet A. 2004;126A:129–40. PubMed PMID: 15057977.

Schanen NC, Kurczynski TW, Brunelle D, Woodcock MM, Dure LS 4th, Percy AK. Neonatal encephalopathy in two boys in families with recurrent Rett syndrome. J Child Neurol. 1998;13:229–31. PubMed PMID: 9620015.

Schönewolf-Greulich B, Tejada MI, Stephens K, Hadzsiev K, Gauthier J, Brøndum-Nielsen K, Pfundt R, Ravn K, Maortua H, Gener B, Martínez-Bouzas C, Piton A, Rouleau G, Clayton-Smith J, Kleefstra T, Bisgaard AM, Tümer Z. The MECP2 variant c.925C>T (p.Arg309Trp) causes intellectual disability in both males and females without classic features of Rett syndrome. Clin Genet. 2016;89:733–8. PubMed PMID: 26936630.

Schüle B, Armstrong DD, Vogel H, Oviedo A, Francke U. Severe congenital encephalopathy caused by MECP2 null mutations in males: central hypoxia and reduced neuronal dendritic structure. Clin Genet. 2008;74:116–26. PubMed PMID: 18477000.

Schwartzman JS, Bernardino A, Nishimura A, Gomes RR, Zatz M. Rett syndrome in a boy with a 47,XXY karyotype confirmed by a rare mutation in the MECP2 gene. Neuropediatrics. 2001;32:162–4. PubMed PMID: 11521215.

Sheikh TI, Ausió J, Faghfoury H, Silver J, Lane JB, Eubanks JH, MacLeod P, Percy AK, Vincent JB. From function to phenotype: impaired DNA binding and clustering correlates with clinical severity in males with missense mutations in MECP2. Sci Rep. 2016;6:38590. PubMed PMID: 27929079.

Sheikh TI, de Paz AM, Akhtar S, Ausió J, Vincent JB. MeCP2_E1 N-terminal modifications affect its degradation rate and are disrupted by the Ala2Val Rett mutation. Hum Mol Genet. 2017;26:4132–41. PubMed PMID: 28973632.

Stallworth JL, Dy ME, Buchanan CB, Chen CF, Scott AE, Glaze DG, Lane JB, Lieberman DN, Oberman LM, Skinner SA, Tierney AE, Cutter GR, Percy AK, Neul JL, Kaufmann WE. Hand stereotypies: lessons from the Rett Syndrome Natural History Study. Neurology. 2019;92:e2594–603. PubMed PMID: 31053667.

Topçu M, Akyerli C, Sayi A, Toruner GA, Kocoglu SR, Cimbis M, Ozcelik T. Somatic mosaicism for a MECP2 mutation associated with classic Rett syndrome in a boy. Eur J Hum Genet. 2002;10:77–81. PubMed PMID: 11896459.

Van Esch H, Bauters M, Ignatius J, Jansen M, Raynaud M, Hollanders K, Lugtenberg D, Bienvenu T, Jensen LR, Gecz J, Moraine C, Marynen P, Fryns JP, Froyen G. Duplication of the MECP2 region is a frequent cause of

severe mental retardation and progressive neurological symptoms in males. Am J Hum Genet. 2005;77:442–53. PubMed PMID: 16080119.

Venâncio M, Santos M, Pereira SA, Maciel P, Saraiva JM. An explanation for another familial case of Rett syndrome: maternal germline mosaicism. Eur J Hum Genet. 2007;15:902–4. PubMed PMID: 17440498.

Venkateswaran S, McMillan HJ, Doja A, Humphreys P. Adolescent onset cognitive regression and neuropsychiatric symptoms associated with the A140V MECP2 mutation. Dev Med Child Neurol. 2014;56:91–4. PubMed PMID: 24328834.

Villard L, Kpebe A, Cardoso C, Chelly PJ, Tardieu PM, Fontes M. Two affected boys in a Rett syndrome family: clinical and molecular findings. Neurology. 2000;55:1188–93. PubMed PMID: 11071498.

Wan M, Lee SS, Zhang X, Houwink-Manville I, Song HR, Amir RE, Budden S, Naidu S, Pereira JL, Lo IF, Zoghbi HY, Schanen NC, Francke U. Rett syndrome and beyond: recurrent spontaneous and familial MECP2 mutations at CpG hotspots. Am J Hum Genet. 1999;65:1520–9. PubMed PMID: 10577905.

Weaving LS, Williamson SL, Bennetts B, Davis M, Ellaway CJ, Leonard H, Thong MK, Delatycki M, Thompson EM, Laing N, Christodoulou J. Effects of MECP2 mutation type, location and X-inactivation in modulating Rett syndrome phenotype. Am J Med Genet A. 2003;118A:103–14. PubMed PMID: 12655490.

Winnepenninckx B, Errijgers V, Hayez-Delatte F, Reyniers E, Frank Kooy R. Identification of a family with nonspecific mental retardation (MRX79) with the A140V mutation in the MECP2 gene: is there a need for routine screening? Hum Mutat. 2002;20:249–52. PubMed PMID: 12325019.

Young JI, Hong EP, Castle JC, Crespo-Barreto J, Bowman AB, Rose MF, Kang D, Richman R, Johnson JM, Berget S, Zoghbi HY. Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proc Natl Acad Sci USA. 2005;102:17551–8. PubMed PMID: 16251272.

Zahorakova D, Rosipal R, Hadac J, Zumrova A, Bzduch V, Misovicova N, Baxova A, Zeman J, Martasek P. Mutation analysis of the MECP2 gene in patients of Slavic origin with Rett syndrome: novel mutations and polymorphisms. J Hum Genet. 2007;52:342–8. PubMed PMID: 17387578.

Zeev BB, Yaron Y, Schanen NC, Wolf H, Brandt N, Ginot N, Shomrat R, Orr-Urtreger A. Rett syndrome: clinical manifestations in males with MECP2 mutations. J Child Neurol. 2002;17:20–4. PubMed PMID: 11913564.

Zhang Q, Yang X, Wang J, Li J, Wu Q, Wen Y, Zhao Y, Zhang X, Yao H, Wu X, Yu S, Wei L, Bao X. Genomic mosaicism in the pathogenesis and inheritance of a Rett syndrome cohort. Genet Med. 2019;21:1330–8. PubMed PMID: 30405208.

Chapter Notes

Author History

Vicky L Brandt; Baylor College of Medicine (2000-2004)
John Christodoulou, MBBS, PhD, FRACP, FRCPA, FHGSA (2006-present)
Gladys Ho, MSc; Children's Hospital at Westmead, Sydney (2009-2019)
Simranpreet Kaur, MSci, MPhil (2019-present)
Huda Y Zoghbi, MD; Baylor College of Medicine (2004-2006)

Revision History

- 19 September 2019 (bp) Comprehensive update posted live
- 28 June 2012 (me) Comprehensive update posted live
- 2 April 2009 (me) Comprehensive update posted live
- 25 January 2008 (cd) Revision: *MECP2* duplication syndrome added to Genetically Related Disorders
- 15 August 2006 (me) Comprehensive update posted live

- 11 February 2004 (me) Comprehensive update posted live
- 3 October 2001 (me) Review posted live
- September 2000 (vb) Original submission

License

Appendix H

Supplementary information for Chapter 7

Detailed clinical report

Individual 1

The proband was born at 38 weeks gestation by normal delivery. She required oxygen for a short period after birth (10-15 minutes), and was observed in Special Care Nursery overnight. She had bilateral cephalohaematoma and had jaundice requiring phototherapy for 2-3 days. Her birth weight was 3.5kg (74th percentile), length 53cm (98th percentile), and head circumference 34 cm (51st percentile). Her parents first became concerned about her development at approximately six months of age when she was not able to roll. She sat unsupported at 14 months, pull to stand at 17 months and learnt to walk at 19 months. As an infant she was not interested in playing with toys and did not reach out or hold an object until approximately 12 months, but did enjoy music. She developed pincer grip after her first birthday and learnt to drink from a cup after two years of age. In addition to her delayed development, there were also concerns of developmental regression. She babbled at 4-5 months but when reviewed at 46 months she vocalised and was not babbling. She also learnt to say "wow" but stop doing that after 2 years. She used to pat her abdomen to indicate she was hungry at one year of age but stopped doing so at two years. After two years of age, she developed stereotypic hand movements in the form of hand flapping, clapping, tapping her face and her mouth with her hands, and hitting her head. Her development was assessed using Griffith Mental Developmental Scale at 45.3 months of age which showed severe to profound global developmental delay. She was able to finger-feed but often after one bite drops food to the floor. She drank from a cup which she could put down neatly on a table but dropped it other times. She was not toilet trained and could not indicate if the nappy needed changing. She communicated by pulling her parents by the hand, to for example, the fridge when hungry. Her vocalisation was limited, consisting of repetitive words such as "ticker, ticker". Words such as "mama" and "papa" were not used in context. She did not respond to her name when being called, although audiology assessment was normal. She had no difficulties chewing and swallowing most food, but meat needed to be cut into small pieces. She walked independently

with a broad-based gait. She walked upstairs holding an adult's hand and placing both feet on the same step, but unable to navigate stairs independently. She was not interested in rolling, throwing or catching a ball. Her finger movements were very agile and that she was able to pick up small items. Her play skills were limited, although she enjoyed her music box. She enjoyed ripping pieces of paper and flicking her fingers in front of her eyes. She had bruxism, occasional constipation, and a disturbed sleep pattern with nocturnal crying.

The Griffith Mental Developmental Assessment was repeated at 5 years and 10 months of age. It was noted at this stage that she had some interval developmental gains. She walked short distances, able to climb stairs holding the rail, and sometimes run in the playground. She was able to feed herself when hungry and can take the top of yoghurt containers. She fed herself predominantly using fingers, but was able to use spoon. She was also able to make choices between foods that she liked. She mouthed a lot of objects such as shoes, leaves, and dirt in her mouth. Her communication was limited and had not developed spoken words. She moved towards an adult when distressed but was not able to share happiness. She cried when angry or frustrated. She was able told a pen and may make some a scribble. This assessment result was consistent with previous findings of severe/profound global developmental disability.

With regard to her medical history, there were concerns of seizures after 5 years of age when she developed episodes where she shook her head and held her hands to the side. There were also reports of staring episodes. EEG performed at 7 years of age did not show definite focal features or epileptiform discharges. Most recently, she had two episodes of generalised tonic-clonic seizures at 18 years 7 months and at 18 years 11 months of age. EEG performed at 19 years and 7 months of age did not show epileptiform activity, however, there were diffused cerebral dysfunction. She was started on Sodium Valproate in increasing doses with a target of 400 mg BD. She was otherwise in good health and occasionally had constipation which was managed with increased fluid intake and fibre.

She was last reviewed at 19 years and 8 months of age. She was described as a fussy eater and had lost some weight. She had been diagnosed with Helicobacter pylori enteritis but did not have any difficulty chewing, swallowing, or any history to suggest aspiration. She had good hand function, was able to finger feed, use a fork to feed herself noodles, and use a spoon for small pieces of food. She was also able to use an iPad. Her constipation had been well managed

with increased fluid, fibre and physical activities. She remained active with independent mobility and was able to walk all day without fatigue or assistance. She had also reached her adult height when reviewed. Her height was 181.9cm (99th centile, Z score of 2.89), her weight was 58.1kg (53rd centile, Z score of 0.08), and the head circumference was 57cm (<1 percentile). She was alert and cooperative throughout assessment. She did not vocalise. Cardiovascular, respiratory and abdominal examinations were normal. She had normal muscle tone and no joint contractures. She was able to walk unassisted with a steady gait. She held a toy with both hands in the midline and constantly played and twisted it. She was able to open a zippered bag and remove objects from the bag. She had truncal rocking in the seated position and tended to hunch forward. Blood test collected after her assessment showed low ferritin level which reflected her difficulties with food choices. She was commenced on iron supplementation. An ECG did not show cardiac conduction abnormalities.

Her genetic investigations included negative *MECP2* screening (exon 1-4), normal MLPA testing, NGS for Epileptic encephalopathy gene panel (TruSight One panel) identified a heterozygous missense variant of uncertain significance in Chromodomain-helicase-DNA-binding protein 2 [*CHD2*; chr15:93558040C>A; NM_001271.3:C.4807C>A; p.(Pro1603Thr)]. Various *in silico* tools consistently predicted this variant to be benign [Mutationtaster: polymorphism (p = 0.693), PolyPhen-2: benign (p=0.025), SIFT: tolerated (score = 0.756; cutoff = 0.05), Grantham score = 38]. Other metabolic investigations were normal including serum transferrin isoforms, plasma very long chain fatty acids, plasma homocysteine, urine amino and organic acids, urine purines and pyrimidines, urine glycosaminoglycans, and white blood cell lysosomal enzymes.

Individual 2

This female, who was the oldest of three daughters, was born at term after an unremarkable pregnancy apart from mild pre-eclampsia to her non-consanguineous parents.

At birth, she had normal Apgar scores and there were no early concerns. She was late to roll and sit and had begun physiotherapy at 4.5 months of age. She was noticed to have global developmental delay from around seven months of age. She sat at 14 months, crawled at 18

months and walked at 2.5 years of age. She had some words but very limited functional speech, but subsequently she had lost even those words as well. However, at 23 years of age, she went through an intense speech therapy, as a consequence of which she had gained several words.

She exhibited slow regression of gross and fine motor skills but did not show any phase of stabilization or recovery. Her physical function declined and she developed an abnormal gait. At 14 years of age she could walk 5 kms but at earlier age she could only walk around 200 meters with less stamina and lower stability. At 23 years and 10 months of age, she was observed to fatigue very easily. She had diminished hand function with age. She used a pincer grip in the past but more recently she was found to rake. She was able to use cutlery until 21 years of age; however she had reduced control of cutlery with age, and at 24 years and 11 months of age, she usually use hands for feeding or was fed by her family. She had severe behavior problems involving self-harm (such as slapping on face repeatedly) when annoyed that was treated with a number of psychoactive medications. She also showed marked involuntary head writhing movements but had no seizures. She had an impaired sleep pattern and woke several times every night. She also had very significant constipation and her feet often turned blue. When last reviewed, she had a diagnosis of ASD and both her siblings also had mild ASD. During last examination at 24 years and 11 months of age, her weight was 65kg (75^{th} percentile), height of 154 cms (5^{th}-10^{th} percentile) and head circumference of 57 cms (75^{th}-98^{th} percentile).

Her brain MRI showed a normal pituitary gland, multiple foci of non-specific white matter T2 hyper-intensity but no structural brain abnormalities. She had no cardiac abnormalities.

Extensive metabolic investigations were performed with normal results: full blood examination, urea and electrolytes, thyroid function tests, iron indices, acylcarnitine and amino acid profiles in blood, urine metabolic amino and organic acids, serum transferrin isoforms, active B12, urine thin layer chromatography, lysosomal enzyme analysis and homocysteine analysis. Serum Prolactin levels were found to be raised but were believed to be secondary due to risperidone therapy. In addition, raised copper but normal caeruloplasmin was also reported.

Negative genetic tests included *MECP2* testing, MLPA, mitochondrial DNA mutations and *POLG* mutations.

Individual 3

She was born at term after a normal pregnancy and delivery to unrelated Ashkenazi Jewish parents with no family history of development delay. Her birth weight was 3.7 kg and she did not have any early feeding problems. She rolled over at 7 months, sat at 12 months, crawled at 17 months of age and walked with support at 23 months of age. She lost eye contact at an early age, but regained it at 20 months of age. She had acquired a few words at 23 months of age which were later lost. She was able to hold toys and transfer them between hands at 18 months of age, but she lost that ability at 2 years of age.

At 2 years and 5 months of age, she could stand and walk around table. She had no functional hand use and exhibited an increased frequency of hand stereotypies. Along with bruxism, she also developed screaming spells. However, she had no breathing abnormalities, had a normal EEG, and head circumference was noted to be 45 cms (<2^{nd} percentile). By 3.5 years of age, she had lost the ability to walk or crawl and had developed mild intermittent apneic episodes, and had mild constipation. At 7 years of age, no further regression was noticed and head circumference was 47 cms (<<2^{nd} percentile). At 10 years of age, absence seizures with intermittent bilateral spikes on awake EEG were observed that responded to valproic acid. She had very good eye contact and communication using an eye tracker apparatus. At 19 years of age, she was wheelchair bound. She could not sit, stand or walk independently and had developed frequent hand mouthing stereotypies. She did not exhibit any further seizures while on valproic acid. Her head circumference was 49 cms (<<< 2^{nd} percentile) and she had severe constipation, bruxism, good eye contact and no facial dysmorphic features.

Supplementary Results

Individual 1

In our previous studies, we had performed targeted sequencing on DNA from individual 1 using an Agilent SureSelect custom designed NGS targeted sequencing panel containing 136

genes implicated in RTT, RTT-like and other NDDs (Appendix A, List A2). Heterozygous Variant of Uncertain significance (VUS) were identified in two candidate genes: inosine monophosphate dehydrogenase 2 gene (*IMPDH2*) and kinesin family member 23 (*KIF23*), also known as Mitotic Kinesin Like-1 (*MKLP1*).

IMPDH2

The heterozygous variant in *IMPDH2* [chr3:49063786C>A; NM_000884: c.977G>T: p.(Gly326Val)] was found to be absent in gnomAD and various *in silico* tools predicted this variant to be likely pathogenic. This variant was predicted to disrupt the conserved enzymatic site of IMP dehydrogenase / GMP reductase that catalyses the conversion of inosine 5'-phosphate (IMP) to xanthosine 5'-phosphate (XMP). This is the first committed and rate-limiting step in the *de novo* synthesis of guanine nucleotides which is vital for the polymerisation and stability of microtubules. However, chromosome microarray using Illumina CytoSNP 850K Microarray did not found any copy number changes within the *IMPDH2* gene. Moreover, her plasma urate, urine purines and pyrimidines levels were also normal suggesting that it is unlikely that the variant in *IMPDH2* is contributing to her phenotype and that she is the carrier of variant in *IMPDH2* gene.

KIF23/ MKLP1

A heterozygous missense variant [chr15: 69728024C>G; NM_138555:exon12:c.1186C>G: (p.Leu396Val)] identified in individual 1 was predicted to effect the motor domain of *KIF23* that is critical for ATP-dependent movement of the protein. In mature neurons, KIF23 transports the minus-end distal microtubules toward the dendrites during neuronal development, and contributes to the multi-polar orientation of dendritic microtubules, as well as contributing to their short and tapered morphology (Xu, He, Zhang, & Chen, 2006; Yu et al., 2000; Yu, Sharp, Kuriyama, Mallik, & Baas, 1997). In differentiating neurons, the movement of short microtubules into the axons is restricted by KIF23, which acts as an opposing force against the driving force of cytoplasmic dynein. Sanger sequencing confirmed this variant to be absent in case of mother of individual 1. However, we could not determine the segregation in father as DNA sample was not available. This variant was present at minor allele frequency of 2.78e-5 in gnomAD and all the *in-silico* tools (SIFT, polyphen-2, mutation

taster) consistently predicted this variant to be disease causing. A homozygous missense variant in *KIF23* (NM_004856: c.755T>A; p.Leu252His) has been previously reported in siblings with severe microcephaly (Karaca et al., 2015). We then analysed WGS data from individual 1 but no second likely pathogenic variant in *KIF23* was identified. The pathogenicity of this variant was further tested by performing a peripheral neurite tip accumulation assay in differentiating Neuro2a cells. For this assay, we used the full length MKLP1 construct tagged with pmCherry at the C-terminus (pmCherry-C1-MKLP1) (Appendix B, Figure B3), which was a gift from Masanori Mishima (Addgene plasmid#70154; http://n2t.net/addgene:70154; RRID:Addgene_70154) (Joseph et al., 2012) and performed site-directed mutagenesis using QuikChange lightning mutagenesis kit (Agilent technologies) according to the manufacturer's instructions to introduce patient-specific variants into the *KIF23* construct. The mutagenesis primers were designed using QuikChange Primer Design (www.agilent.com/genomics/qcpd; Forward primer: catcaaagtagttcttgaacacatgggttaactttgaatctcgata and reverse primer: tatcgagattcaaagttaacccatgtgttcaagaactactttgatg) and confirmed by Sanger sequencing using a second primer pair (Forward primer: accagagcagaagggaacag and reverse primer: ttcttcataatcttcagccttgg). Neuro2a cells were grown in DMEM media (DMEM with 3.7g/L $NaHCO_3$, 10% (vol/vol) fetal bovine serum (FBS), 1X Pen-Strep) at 37°C in a humidified incubator and 5% CO_2. For the neurite tip accumulation assay, Neuro2a cells were transfected with *MKLP1*/*KIF23* using Lipofectamine 2000. After 24 hours, cells were washed with 1X PBS, fixed in 4% (vol/vol) paraformaldehyde (PFA) in PBS for 20 minutes at room temperature before mounting on glass slides using Prolong Gold with DAPI (Life Technologies). Images were acquired using Zeiss AxioVision fluorescence microscope with 40X objective with zoom factor of two. No difference in translocation efficiency was observed between variant and WT construct (Figure S5).

Table H1: List of the genes involved in the *de novo* deletion at chromosome 9q34.11.

Gene	Gene name	Function	Gene Ontology (GO)- Function	MIM	Disease
SET	SET nuclear proto-oncogene	Inhibits acetylation of nucleosomes by HAT, nucleosome assembly, histone chaperoning	histone binding	600960	Mental retardation, autosomal dominant 58
PKN3	protein kinase N3	Activates PI3K/PTEN signal transduction pathway	transferase activity, transferring phosphorus-containing groups, protein tyrosine kinase activity	610714	NA
ZDHHC12	zinc finger DHHC-type containing 12	Palmitoyltransferase activity	protein-cysteine S-palmitoyltransferase activity, palmitoyltransferase activity	NA	NA
LOC100506100	uncharacterized LOC100506100	Unknown function, belongs to long non-coding RNA class	binding and ubiquitin-protein transferase activity	NA	NA
ZER1	zyg-11 related cell cycle regulator	Encodes a subunit of an E3 ubiquitin ligase complex that may be involved in meiosis.	binding, ubiquitin-protein transferase activity	617764	NA
TBC1D13	TBC1 domain family member 13	Acts as a GTPase-activating protein for RAB35	GTPase activator activity	616218	NA
ENDOG	endonuclease G	Cleaves DNA at double-stranded (DG)n.(DC)n and at single-stranded (DC)n tracts	nucleic acid binding, endonuclease activity	600440	NA
SPOUT1	SPOUT domain containing methyltransferase 1	Associates centrosomes with the poles of the bipolar mitotic spindle during metaphase, helps in chromosome alignment; may promote centrosome maturation probably by recruiting A-kinase anchor protein AKAP9 to centrosomes in early mitosis; Binds specifically to miRNA miR145 hairpin, regulates miR145 expression at a post-transcriptional level	Methyltransferase activity miRNA binding RNA binding	617614	NA
KYAT1	Kynurenine aminotransferase 1	Catalyzes the irreversible transamination of the L-tryptophan metabolite L-kynurenine to form kynurenic acid (KA)	cysteine-S-conjugate beta-lyase activity, glutamine-phenylpyruvate transaminase activity, kynurenine-oxoglutarate transaminase activity, protein homodimerization activity, pyridoxal phosphate binding, transaminase activity	600547	NA

Gene	Gene name	Function	Gene Ontology (GO)- Function	MIM	Disease
LRRC8A	Leucine rich repeat containing 8 VRAC subunit A	Essential component of the volume-regulated anion channel (VRAC) to maintain a constant cell volume in response to extracellular or intracellular osmotic changes	volume-sensitive anion channel activity	608360	Agammaglobulinemia 5
PHYHD1	Phytanoyl-CoA dioxygenase domain containing 1	alpha-ketoglutarate-dependent dioxygenase activity	dioxygenase activity, metal ion binding	NA	NA
DOLK	Dolichol kinase	Catalyzes the CTP-mediated phosphorylation of dolichol, and is involved in the synthesis of Dol-P-Man, which is an essential glycosyl carrier lipid for C- and O-mannosylation, N- and O-linked glycosylation of proteins, and for the biosynthesis of glycosyl phosphatidylinositol anchors in endoplasmic reticulum.	Dolichol kinase activity	610746	Congenital disorder of glycosylation, type IM
NUP188	Nucleoporin 188	Component of the nuclear pore complex	Structural constituent of nuclear pore	615587	NA

Table H2: **List of VUS identified in Individual 2 through WES and their curation.**

Gene	Gene name	Variant	Zygosity	Type	Exon	gnomAD	ClinVar	*In silico*	Segregation
NSD1	Nuclear receptor binding SET Domain Protein 1	chr5:176721240C>T NM_022455.4: c.6871C>T; NP_071900.2: p.(Gln2291*)	Het	Stop gain	Exon 23/23	7.98e-6	Absent	Disease causing	Inherited from father
NSD1	Nuclear receptor binding SET Domain Protein 1	chr5:176721242G>T NM_022455.4: c.6873G>T; NP_071900.2: p.(Gln2291His)	Het	Stop gain	Exon 23/23	7.98e-6	Absent	Polymorphism	Inherited from father
ZNF557	Zinc finger protein 557	chr19:7083170C>G NM_024341.2: c.708C>G; NP_077317.2:p.(Tyr236*)	Het	Stop gain	Exon 8/8	Absent	Absent	Disease causing	Inherited from mother
CNTNAP2	Contactin associated protein 2	chr7:147964217A>G NM_014141.5: c.3474A>G; NP_054860.1:p.(Ile1158Met)	Het	Missense / splice site	Exon 21/24	Absent	Present (1 VUS)	Polymorphism	Inherited from father
CNTNAP4	Contactin associated protein family member 4	Chr16:76389241T>G; NM_138994.3: c.148T>G; NP_620481.2: p.(Ser50Ala)	Het	Missense	Exon 3/23	Absent	Absent	Disease causing	Inherited from mother
SCN9A	Sodium voltage-gated channel alpha subunit 9	chr2:167056258G>A; NM_002977.3: c.4858C>T; p.(Arg1620Cys)	Het	Missense	Exon 27/27	1.77e-5	Absent	Disease causing	Inherited from mother
SYT16	Synaptotagmin 16	chr14:62550972C>T; NM_031914.2: c.1493C>T; p.(Thr498Met)	Het	Missense	Exon 6/6	2.14e-5	Absent	Disease causing	Inherited from father
KLC4	Kinesin light chain 4	chr6:43029152C>T; NM_138343.2: c.79C>T; NP_612352.1: p.(Arg27Trp)	Het	Missense	Exon 2/6	2.91e-4	Absent	Disease causing	Inherited from father

Table H3: Clinical feature comparison between individual 2 and previously published individuals with *SET* variants.

Parameter	Present study (Individual 2)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	DDD study (2017)	DDD study (2017)	DDD study (2017)	Hamdan et al (2014)
Variant	deletion encompassing 3' region of SET (stop codon at 178 amino acid)	c.167_170del; p.Arg57Leufs*10	c.167_170del; p.Arg57Leufs*10	c.283T > G; p.Trp95Gly	c.352C > T; p.His118Tyr	c.417+1G > C; r.spl?	c.689_690dup; p.Gln231Tyrfs*29	c.459_460del; p.Lys154ArgfsTer6	NR	NR	c.699_701delCTT; p.(Tyr233*)
Inheritance	*de novo*	Maternal	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*	*de novo*
Gender	Female	Male	Female	Male	Male	Male	Male	Male	Female	Male	Male
Age	24 years and 4 months	16 years	49 years	4 years	4 years and 3 months	3 years	5 years	NR	NR	NR	12 years
Weight; kg (percentile)	65 (<1st)	69.6	47.25	24.5	16.8	11.1	19.6	NR	NR	NR	21.6 kg (2nd)
Height; cms (percentile)	154 (8th)	170	152	122	98.7	87.6	114	NR	NR	NR	121.3 cm (2nd)
Head circumference; cms (percentile)	57 cms (99th)	57.8	NR	52	49	45.5	52	NR	NR	NR	48 cms (<1st)
Development delay	Yes	Yes	NR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ASD	Yes	No	NR	No	Yes^^	NR		NR	NR	Yes	No
ID	Yes; moderate	Yes; mild	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes; learning disability	Yes; moderate

Parameter	Present study (Individual 2)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	DDD study (2017)	DDD study (2017)	DDD study (2017)	Hamdan et al (2014)
Motor problems	Yes; regression in physical function	Yes; mild incoordination	NR	Yes; delay and frequent falls	Yes; delay; received physiotherapy	Yes; delay	Yes; delay; bilateral fine motor dexterity	No	NR	NR	No
Muscle tone	Normal	NR	NR	Hypotonia (limbs)	Hypotonia (face)		Hypotonia (axial)		Generalized hypotonia	NR	Normal
Brain MRI	Normal	Normal	NR			NR	Normal	NR	NR	NR	NR
Speech delay	Yes; severe^	Yes; able to read short text	Past history	Yes; severe; <20 words	Yes; severe; 40 words	Yes; severe	Yes; severe	NR	NR	NR	Yes; severe; 20 words
Seizures	No	No	No	No	No	No	Yes; two epileptic seizures (4 years and 5 years of age); on antiepileptic medication	NR	NR	NR	NR
Behaviour problems	Yes; self-harm	No	No	Happy disposition	Repellent, fearful, little eye contact	Social, kind, naughty	None	NR	NR	Yes	Yes; ADD

Parameter	Present study (Individual 2)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	DDD study (2017)	DDD study (2017)	DDD study (2017)	Hamdan et al (2014)
Dysmorphic features											
Face		Mild facial asymmetry	None	Mild facial asymmetry due to residual plagiocephaly pointed chin	Hypertelorism, downslanted palpebral fissures, broad nasal bridge, large posteriorly rotated ears, upturned lobes, hypotonic face, small chin	Periorbital fullness, slight hypertelorism, thick lips, broad nasal tip	Mildly asymmetric ears; right ear smaller than the left	Broad forehead hypertelorism thick lower lip vermilion	Depressed nasal bridge Synophrys		NR
Mouth		Submucosal cleft palate, high palate, dental crowding	NR	Wide mouth, occasional tongue protrusion	NR	NR	Cleft soft palate	NR	NR	NR	NR
Hands		Digital fat pads	Dilated digital veins	NR	Unilateral simian crease	NR	NR	NR	NR	NR	NR

Parameter	Present study (Individual 2)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	Stevens et al (2018)	DDD study (2017)	DDD study (2017)	DDD study (2017)	Hamdan et al (2014)
Others		Hypoplastic fifth toenails; mild starbismus; tendency for otitis media; repair of inguinal hernia; velo-pharyngeal insufficiency	Pectus carinatum, possible mild scoliosis	Plagiocephaly, generalized hypermobility, limb hypotonia, more apparent in the lower limbs,with brisk reflexes	Mild neonatal feeding problems, inverted nipples	NR	Growth hormone deficiency, 2 café-au-lait macules	Generalized joint laxity Strabismus	Hypertrichosis	Generalized joint laxity 2–3 toe syndactyly strabismus	Left renal agenesis

```
H sapiens         VLDNFSDIVEGVDFDKDCEDPEYKPLQGPPKDQDDEGDPLMMMDEEISVIDGDEAQVTQQ     60
M musculus        VLDNFSDLVEGIDFDKDCEDPEYKPLQGPPKDPDDEGDPLMMMDEEISVIDGEEAQVTQQ     60
R norvegicus      VLDNFSDLVEGIDFDKDCEDPEYKPLQGPPKDPDDEGDPLMMMDEEISVIDGDEAPVTQQ     60
                  *******:***:******************** ********************:** ****

H sapiens         PGHLFWPPGSALTARLRRLVTAYQRSYKREQMKIEAAERGDRRRRRCEAAFKLKEIARRE     120
M musculus        PGHLFWPPGSALTARLRRLVTAYQRSYKREQMKMEAAERGDRRRRRCEAAFKLKEIARRE     120
R norvegicus      PGHLFWPPGSALTARLRRLVTAYQRSYKREQMKIEAAERGDRRRRRCEAAFKLKEIARRE     120
                  ********************************:*****************************

H sapiens         KQQRWTRREQTDFYRVVSTFGVEYDPDTMQFHWDRFRTFARLDKKTDESLTKYFHGFVAM     180
M musculus        KQQRWTRREQTDFYRVVSTFGVEYDPDNMQFHWDRFRTFARLDKKTDESLTKYFHGFVAM     180
R norvegicus      KQQRWTRREQTDFYRVVSTFGVEYDPDNMQFHWDRFRTFARLDKKTDESLTKYFHGFVAM     180
                  ****************************.*****************************

H sapiens         CRQVCRLPPAAGDEPPDPNLFIEPITEERASRTLYRIELLRRLREQVLCHPLLEDRLALC     240
M musculus        CRQVCRLPPAAGDEPPDPNLFIEPITEERASRTLYRIELLRRLREQVLCHPLLEDRLALC     240
R norvegicus      CRQVCRLPPAAGDEPPDPNLFIEPITEERASRTLYRIELLRRLREQVLCHPLLEDRLALC     240
                  ************************************************************

H sapiens         QPPGPELPKWWEPVRHDGELLRGAARHGVSQTDCNIMQDPDFSFLAARMNYMQNHQAGAP     300
M musculus        QPPGLELPKWWEPVRHDGELLRGAARHGVSQTDCNIMQDPDFSFLAARMNYMQNHQAGAS     300
R norvegicus      QPPGLELPKWWEPVRHDGELLRGAARHGVSQTDCNIMQDPDFSFLAARMNYMQNHQAGAS     300
                  **** ************************************************** *****

H sapiens         APSLSRCSTPLLHQQYTSRTASPLPLRPDAPVEKSPEETATQVPSLESLTLKLEHEVVAR     360
M musculus        AASLSRCSTPLLHQQCTSRTASPSPLRPDAPVEKSPEESTVQVPNLESLTLKLEDEVVAR     360
R norvegicus      AASLSRCSTPLLHQQCTSRTASPSPLRPDVPAEKSPEENAVQVPSLDSLTLKLEDEVVAR     360
                  * ************* ******* *****.*.******.:.***.*:*******.*****

H sapiens         SRPTPQDYEMRVSPSDTTPLVSRSVPPVKLEDEDDSDSELDLSKLSPSSSSSSSSSSSS     420
M musculus        SRLTSQDYEVRVGSSDTAPL-SRSVPPVKLEDEDDSDSELDLSKLSPSSSSSSSSSSSS     419
R norvegicus      SRLTPQDYEIRVASSDTAPL-SRSVPPVKLEDDDDSDSELDLSKLSPSSSSSSSSSSSS     419
                  ** * ****:**. ***:** ***********:****************************

H sapiens         STDESEDEKEEKLT-DQSRSKLYDEESLLSLTMSQDGFPNEDGEQMTPELLLLQERQRAS     479
M musculus        STDESEDEKEEKLTADRSRPKLYDEESLLSLTMSQDGFPNEDGEQMTPELLLLQERQRAS     479
R norvegicus      SSDESEDEKEEKLTADRSRPKLYDEESLLSLTMSQDGFPNEDGEQMTPELLLLQERQRAS     479
                  *:************ *:** ************************************

H sapiens         EWPKDRVLINRIDLVCQAVLSGKWPSSRRSQEMVTGGILGPGNHLLDSPSLTPGEYGDSP     539
M musculus        EWPKDRVLINRIDLVCQAVLSGKWPSNRRSQEVTAGGILGPGNHLLDSPSLTPGEDGDSP     539
R norvegicus      EWPKDRVLINRIDLVCQAVLSGKWPSNRRSQEMTTGGILGPGNHLLDSPSLTPGEYGDSP     539
                  ****************************.*****:.:****************** ****

H sapiens         VPTPRSSSAASMAEEEASAVSTAAAQFTKLRRGMDEKEFTVQIKDEEGLKLTFQKHKLMA     599
M musculus        VPTPRSGSAASMAEEEASAVTTAAAQFTKLRRGMDEKEFTVQIKDEEGLKLTFQKHRLMA     599
R norvegicus      VPTPRSSSAASMVEEEASAVTTAAAQFTKLRRGMDEKEFTVQIKDEEGLKLTFQKHRLMA     599
                  ******.*****.*******:*************************************:***

H sapiens         NGVMGDGHPLFHKKKGNRKKLVELEVECMEEPNHLDVDLETRIPVINKVDGTLLVGEDAP     659
M musculus        NGVMGDGHPLFHKKKGNRKKLVELEVECMEEPNHLDLDLETRIPVINKVDGTLLVGDEAP     659
R norvegicus      NGVMGDGHPLFHKKKGNRKKLVELEVECMEEPNHLDVDLETRIPVINKVDGTLLVGDEAP     659
                  ************************************:******************::**

H sapiens         RRAELEMWLQGHPEFAVDPRFLAYMEDRRKQKWQRCKKNNKAELNCLGMEPVQTANSRNG     719
M musculus        RRAELEMWLQGHPEFAVDPRFLAYMEERRKQKWQRCKKNNKAELNCLGMEPVQPANSRNG     719
R norvegicus      RRAELDMWLQGHPEFAVDPRFLAYMEERRKQKWQRCKKNNKTELNCLGMEPVQPANSRNG     719
                  *****:********************:**************:************ ******

H sapiens         KKGHHTETVFNRVLPGPIAPESSKKRARRMRPDLSKMMALMQGGSTGSLSLHNTFQHSSS     779
M musculus        KKGHYAETAFNRVLPGPVAPENSKKRVRRTRPDLSKMMALMQGGSTGSLSLHNTFQHSSS     779
R norvegicus      KKGHYAETAFNRVLPGPIAPENSKKRVRRTRPDLSKMMALMQGGSTGSLSLHNTFQHSSS     779
                  ****::**.********:***.****.** ******************************

H sapiens         GLQSVSSLGHSSATSASLPFMPFVMGGAPSSPHVDSSTMLHHHHHHPHPHHHHHHHPGLR     839
M musculus        NLQSVSSLGHSSTTSASLPFMPFVMGA-AAPPHVDSSTMLHHHHHHPHPHHHHHHHPGLR     838
R norvegicus      NLQSVSSLGHSSATSASLPFMPFVMGGAAAPPHVDSSTMLHHHHHHPHPHHHHHHHPGLR     839
                  .************:*************.  : ******************************

H sapiens         APGYPSSPVTTASGTTLRLPPLQPEEDDDEDEEDDDDLSQGYDSSERDFSLIDDPMMPAN     899
M musculus        TTGYPSSPATTTSGTALRLPTLQPEDDDEEEDEEDDDLSQGYDSSERDFSLIDDPMMPAN     898
R norvegicus      TTGYPSSPATTTSGTALRLPTLQHEDDDEEEDEDDDDLSQGYDSSERDFSLIDDPMMPAN     899
                  : ******.**:***:**** ** *:**:*::*:**************************

H sapiens         SDSSEDADD  908
M musculus        SDSSEDADD  907
R norvegicus      SDSSDDADD  908
                  ****:****
```

Figure H1: **Multiple sequence alignment showing evolutionary conservation of region affected by CHD8 variants in individual 1**.

The sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_001164100.1), *M musculus* (NP_963999.2), *R norvegicus* (NP_001334590.1). * (asterisk) denotes fully conserved residue, : (colon) denotes conservation between residues with strongly similar properties and a . (period) indicates conservation between residues with weakly similar properties.

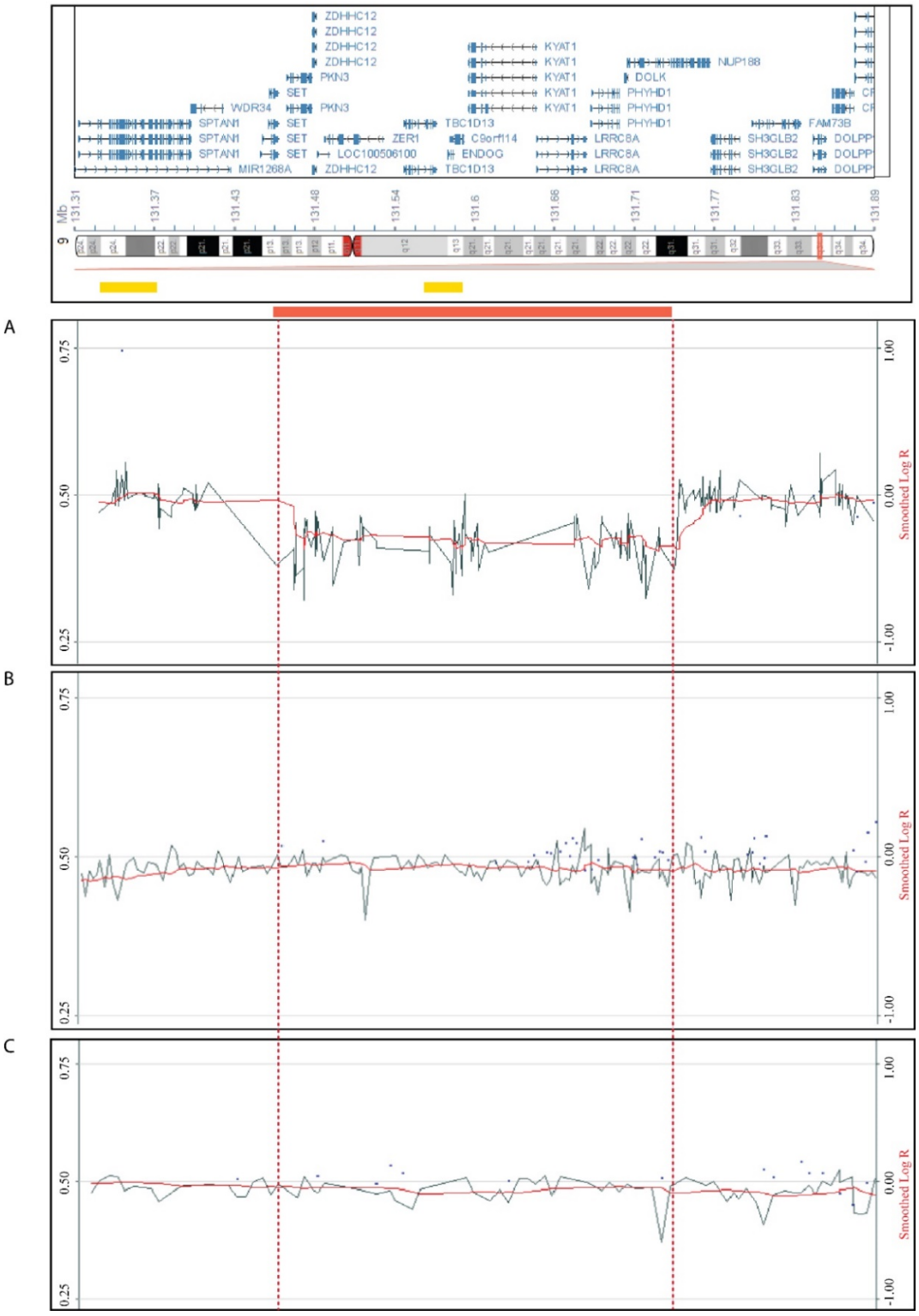


Figure H2: **Illumina Karyostudio microarray profiles for patient 2 and her family.**

The schematic showing the deleted region in the proband (A:horizontal red line) and corresponding normal region in maternal (B) and paternal samples (C) (vertical serrated red lines). blue dots=B allele ratio, red line=smoothed LogR, black line=raw LogR.

```
H sapiens          MAPKRQSPLPPQKKKPRPP-PALGPEETSASAGLPKKGEKEQQEAIEHIDEVQNEIDRLN        59
M musculus         MAPKRQSAILPQPKKPRPA-AAPKLEDKSASPGLPK-GEKEQQEAIEHIDEVQNEIDRLN        58
R norvegicus       MAPKRQSAILPQPKKPRPV-AAPKLEDKSASPGLPK-GEKEQQEAIEHIDEVQNEIDRLN        58
G gallus           --------------MSAPAAKVSKKELNSNHDGADETSEKEQQEAIEHIDEVQNEIDRLN        46
P troglodytes      MAPKRQSPLPPQKKKPRPP-PALGPEETSASAGLPKKGEKEQQEAIEHIDEVQNEIDRLN        59
S scrofa           --------------MSAPAAKVSKKELNSNHDGADETSEKEQQEAIEHIDEVQNEIDRLN        46
E caballus         MAPKRQNSLPPQTKKPRLP-PAPRPEETSKTSVNLPQGEKEQQEAIEHIDEVQNEIDRLN        59
                                    .   * .*        .**********************

H sapiens          EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        119
M musculus         EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        118
R norvegicus       EQASEEILKVEQKYNKLRQPFFQKRSELITKIPNFWVTTFVNHPQVSALLGEEDEEALHY        118
G gallus           EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        106
P troglodytes      EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        119
S scrofa           EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        106
E caballus         EQASEEILKVEQKYNKLRQPFFQKRSELIAKIPNFWVTTFVNHPQVSALLGEEDEEALHY        119
                   ****************************:******************************

H sapiens          LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        179
M musculus         LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        178
R norvegicus       LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        178
G gallus           LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        166
P troglodytes      LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        179
S scrofa           LTRVEVTEFEDIKSGYRIDFYFDENPYFENKVLSKEFHLNESGDPSSKSTEIKWKSGKDL        166
E caballus         LTRVEVTEFEDIKSGYRIDFYFDENPYFENKILSKEFHLNESGDPSSKSTEIKWKSGKDL        179
                   ****************************:******************************

H sapiens          TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        239
M musculus         TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        238
R norvegicus       TKCSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        238
G gallus           TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        226
P troglodytes      TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        239
S scrofa           TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        226
E caballus         TKRSSQTQNKASRKRQHEEPESFFTWFTDHSDAGADELGEVIKDDIWPNPLQYYLVPDMD        239
                  ** ************************************************************

H sapiens          DEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD                290
M musculus         DEEGEAEDDDDDDEEEEGLEDIDEEGDEDEGEEDDDEDEGEEGEEDEGEDD                289
R norvegicus       DEEGEAEDDDDDDEEEEGLEDIDEEGDEDEGEEDNDEDEGEEGEEDEGEDD                289
G gallus           DEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD                277
P troglodytes      DEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD                290
S scrofa           DEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD                277
E caballus         DEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD                290
                   *****.*:**************************:*:**************
```

Figure H3: Multiple sequence alignment showing evolutionary conservation of SET.

The sequences were aligned using Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) using the RefSeqs of *H sapiens* (NP_001116293.1), *M musculus* (NP_076360.1), *R norvegicus* (NP_001012522.1), X tropicalis (NP_989041.1), *G gallus* (NP_001025862.2), *P troglodytes* (XP_024201677.1), *S scrofa* (NP_001231019.1), *E caballus* (XP_023484824.1).

```
H sapiens          MAAAAGGPCVRSSRELWTILLGRSALRELSQIEAELNKHWRRLLEGLSYYKPPSPSSAEK     60
M musculus         MAAAAGGPCVRSSRELWTILLGRSALRELNQIEAELNKYWQRLLEGLSYYKPPSPSSAER     60
R norvegicus       MAAAAGGPCVRSSRELWTILLGRSALRELNQIEAELNKHWQRLLEGLSYYRPPSPSSAER     60
G gallus           --MAAGGLCARSSRELWTILLGRSALREPAQIASELNKHWQRLLEGLSYYKPPSTTSAEK     58
P troglodytes      MAAAAGGPCVRSSRELWTILLGRSALRELSQIEAELNKHWRRLLEGLSYYKPPSPSSAEK     60
B taurus           MAAAAGGQCVRSSRELWTILLGRSALRELNQIEAELNKHWQRLLEGLSYYKPPSPSSAEK     60
                      **** *.****************  ** :****:*:*********:*** :***:

H sapiens          VKANKDVASPLKELGLRISKFLGLDEEQSVQLLQCYLQEDYRGTRDSVKTVLQDERQSQA     120
M musculus         VKANKDVASPLKELGLRVSKFLGLDEEQSVQLLQCYLQEDYRGTRDSLKTVLQDERQSQA     120
R norvegicus       VKANKDVASPLKELGLRVSKFLGLDEEQSVQLLQCYLQEAYRGTRDSLKTVLQDERQSQA     120
G gallus           IKADKDVAAPLKELGLRVSKFLGLDEEQSV-LLQCYLQEDYRGTRDSLKTVLQDERQSQA     117
P troglodytes      VKANKDVASPLKELGLRISKFLGLDEEQSVQLLQCYLQEDYRGTRDSVKTVLQDERQSQA     120
B taurus           VKASKDVASPLKELGLRISKFLGLDEEQSVQLLQCYLQEDYRGTRDSLKTVLQDERQSQA     120
                   :**.****:********:************ ******** *******:************

H sapiens          LILKIADYYYEERTCILRCVLHLLTYFQDERHPYRVEYADCVDKLEKELVSKYRQQFEEL     180
M musculus         LTLKIADYYYEERTCILRCVLHLLTYFQDERHPYRAEYADCVDKLEKELVLKYRQQFEEL     180
R norvegicus       LTLKIADYYYEERTCILRCVLHLLTYFQDERHPYRAEYADCVDKLEKELVLKYRQQFEEL     180
G gallus           LMLKLADYYYEERICILRCVLHLLTYFQDERHPYTAQYFQCVEKLDKELVPNYRKQFEEL     177
P troglodytes      LILKIADYYYEERTCILRCVLHLLTYFQDERHPYRVEYADCVDKLEKELVSKYRQQFEEL     180
B taurus           LILKIADYYYEERTCILRCVLHLLTYFQDERHPYRVEYADCVDKLEKELVTKYRQQFEEL     180
                   * **:********* ******************** .:* :**:**:**** :**:*****

H sapiens          YKTEAPTWETHGNLMTERQVSRWFVQCLREQSMLLEIIFLYYAYFEMAPSDLLVLTKMFK     240
M musculus         YRTEAPTWETHGNLMTERQVSRWLVQCLREQSMLLEIIFLYYAYFEMAPSDLLVLTKMFK     240
R norvegicus       YRTEAPTWETHGNLMTERQVSRWFVQCLREQSMLLEIIFLYYAYFEMAPSDLLVLTKMFK     240
G gallus           YKAEAPTWETHGNLMTERQVSRWFVQCLREQSMLLEVIFLYYAYFEMAPDELLAFAKLFK     237
P troglodytes      YKTEAPTWETHGNLMTERQVSRWFVQCLREQSMLLEIIFLYYAYFEMAPSDLFVLTKMFK     240
B taurus           YKTEAPTWETHGNLMTERQVSRWFVQCLREQSMLLEIIFLYYAYFEMAPSDLLVLTKMFK     240
                   *::*********************:************:************.:*:.::*:**

H sapiens          EQGFGSRQTNRHLVDETMDPFVDRIGYFSALILVEGMDIESLHKCALDDRRELHQFAQDG     300
M musculus         EQGFGSRQTSRHLVGGTMDPFVDRIGYFSALILVEGMDIESLHKYALDDRRELHQFAQDG     300
R norvegicus       EQGFGSRQTNRHLVDETMDPFVDRIGYFSALILVEGMDIESLHKYALDDRRELHQFAQDG     300
G gallus           EQGFGTRQTNRHLVDESMDHLVDRIGYFSALILVEGMEIDSLHERALDDRTEEHKLANNS     297
P troglodytes      EQGFGSRQTNRHLVDETMDPFVDRIGYFSALILVEGMDIESLHKCALDDRRELHQFAQDG     300
B taurus           EQGFGSRQTNRHLVDETMDPFVDRIGYFSALILVEGMDIESLHKCALDDRRELHQFAQDG     300
                   *****:***.****. :** :*****************:*:***: ***** * *::*::.

H sapiens          LICQDMDCLMLTFGDIPHHAPVLLAWALLRHTLNPEETSSVVRKIGGTAIQLNVFQYLTR     360
M musculus         LICQDMDRAMLTLGDIPHHAPVLLAWALLRHTLSPEETSSVVRKIGGTAIQLNVFQYLTR     360
R norvegicus       LICQDMDRIMLTLGDIPHHAPVLLAWALLRHTLNPEETSSVVRKIGGTAIQLNVFQYLTR     360
G gallus           HIHQEMDNQLLQLGDVSHHAPVLLAWALLRHTINPEESANVIRRMGSTAIQLNVFQYLTK     357
P troglodytes      LICQDMDCLMLTFGDIPHHAPVLLAWALLRHTLNPEETSSVVRKIGGTAIQLNVFQYLTR     360
B taurus           LICQDMDRLMLTFGDIPHHAPVLLAWALLRHTLNPEETSSVVRKIGGTAIQLNVFQYLTR     360
                    * *:**  :* :**: ***************:.***::.*:*::*.************:

H sapiens          LLQSLASGGNDCTTSTACMCVYGLLSFVLTSLELHTLGNQQDIIDTACEVLADPSLPELF     420
M musculus         LLRSLASGGNDCTTSTACMCVYGLLSFALTSLELHTLGNQQDVIDTACEVLADPSLPELF     420
R norvegicus       LLRSLASGGNDCTTSTACMCVYGLLSFALTSLELHTLGNQQDVIDTACEVLADPSLPELF     420
G gallus           LLKSLSSGGKNCTASTACMCIYGLLSFVLTSLELHTLGNQQDIIDAACEVLAASSIPQLF     417
P troglodytes      LLQSLASGGNDCTTSTACMCVYGLLSFVLTSLELHTLGNQQDIIDTACEVLADPSLPELF     420
B taurus           LLRSLASGGNDCTTSTACMCVYGLLSFVLTSLELHTLGNQQDVIDTACEVLADPSLPELF     420
                   **:**:***::**:******:******.**************:**:******  *:*:**

H sapiens          WGTEPTSGLGIILDSVCGMFPHLLSPLLQLLRALVSGKSTAKKVYSFLDKMSFYNELYKH     480
M musculus         WGTEPTSGLGIILDSVCGMFPHLLSPLLQLLRALVSGKSTAKKVYSFLDKMSFYNELHKH     480
R norvegicus       WGTEPTSGLGIILDSVCGMFPHLLSPLLQLLRALVSGKSTAKKVYSFLDKMSFYNELHKH     480
G gallus           WKTEPTAGLGIILDSVCGMFPHLLTPLLQLLQALVSDKSTAKKVYSFLDKMSFYIELYKH     477
P troglodytes      WGTEPTSGLGIILDSVCGMFPHLLSPLLQLLRALVSGKSTAKKVYSFLDKMSFYNELYKH     480
B taurus           WGTEPTSGLGIILDSVCGMFPHLLSPLLQLLRALVSGKSTAKKVYSFLDKMSFYNELYKH     480
                   * ****:******************:******:****.**************** **:**

H sapiens          KPHDVISHEDGTLWRRQTPKLLYPLG-GQTNLRIPQGTVGQVMLDDRAYLVRWEYSYSSW     539
M musculus         KPHDVLSHEDGTLWRRQTPKLLYPLG-GQTNLRIPQGTVGQVMLDDRAYLVRWEYSYSSW     539
R norvegicus       KPHDVLSHEDGTLWRRQTPKLLYPLG-GQTNLRIPQGTVGQVMLDDRAYLVRWEYSYSSW     539
G gallus           KPPDVISHEDGTLWRRQAPKLLYPLGLGQTNLRIPQGTVGQVMLDDRAYLVRWEYSYSSW     537
P troglodytes      KPHDVISHEDGTLWRRQTPKLLYPLG-GQTNLRIPQGTVGQVMLDDRAYLVRWEYSYSSW     539
```

```
H sapiens          LTTVISPPVDVIASCVNCLTVLAARNPAKVWTDLRHTGFLPFVAHPVSSLSQMISAEGMN     659
M musculus         LTTVISPPVNVIASCVNCLTVLAARNPAKVWTDLRHTGFLPFVAHPVSNMTQMISAEGMN     659
R norvegicus       LTTVISPPVNVIASCVNCLTVLAARNPAKVWTDLRHTGFLPFVAHPVSNMTQMISAEGMN     659
G gallus           LTTVISPPMDVIASCVNCLTVLAARMPAKVWTDLHHTGFLPFVANPVSSMNHMISAEGMN     657
P troglodytes      LTTVISPPVDVIASCVNCLTVLAARNPAKVWTDLRHTGFLPFVAHPVSSLSQMISAEGMN     659
B taurus           LTTVISPPVDVIASCVNCLTVLAARNPAKVWTDLLHTGFLPFVAHQVSSMSQMISAEGMN     659
                   ********::*************** ******** *********: **.:.:******** 
H sapiens          AGGYGNLLMNSEQPQGEYGVTIAFLRLITTLVKGQLGSTQSQGLVPCVMFVLKEMLPSYH     719
M musculus         AGGYGSLLMNSEQPQGEYGVTIAFLRLVTTLVKGQLGSTQSQGLVPCVMFVLKEMLPSYH     719
R norvegicus       AGGYGSLLMNSEQPQGEYGVTVAFLRLVTTLVKGQLGSTQSQGLVPCVMFVLKEVLPSYH     719
G gallus           AGGYGNLLMSIEQPQGEYSVTISFLNLITTLVRGQLGSTQSQGLVPCIVFVLKEMLPNYH     717
P troglodytes      AGGYGNLLMNSEQPQGEYGVTIAFLRLITTLVKGQLGSTQSQGLVPCVMFVLKEMLPSYH     719
B taurus           AGGYGNLLMNSEQPQGEYGVTVAFLRLITTLVKGQLGSTQSQGLVPCVMFVLKEMLPSYH     719
                   *****.***. *******.**::**.*:****:**************::*****:**.** 
H sapiens          KWRYNSHGVREQIGCLILELIHAILNLCHETDLHSSHTPSLQFLCICSLAYTEAGQTVIN     779
M musculus         KWRYNSHGVRELIGCLILELIHAILNLCQETELHSSHTPSLPSLCICSLAYTEAGQTVIS     779
R norvegicus       KWRYSSHGVRELIGCLILELIHAILNLCQETELHSSHTPSLPSLCICSLAYTEAGQTVIS     779
G gallus           KWRYNSHGVREKIGCLILQLIHAILNLCPEMDPRSSNAPSLQSLCIFSLANTEAGQAVIN     777
P troglodytes      KWRYNSHGVREQIGCLILELIHAILNLCHETDLHSSHTPSLQSLCICSLAYTEAGQTVIN     779
B taurus           KWRYNSHGVREQIGCLILELIHSILNLCHETDLHSSHTPSLQSLCICSLAYTEAGQTVIN     779
                   ****.****** ******:***:***** * : :**::***  *** *** *****:**. 
H sapiens          IMGIGVDTIDMVMAAQPR---SDGAEGQGQGQLLIKTVKLAFSVTNNVIRLKPPSNVVSP     836
M musculus         IMGIGVDTIDMVMAAQPR---SDGPEGQGQGQLLIKTVKLAFSVTNNVIRLKPPSNVVSP     836
R norvegicus       IMGIGVDTIDMVMAAQPR---SDGPEGQGQGQLLIKTVKLAFSVTNNVIRLKPPSNVVSP     836
G gallus           IMGIGVDTIDMVMASQS-----SSDESQGQGQLLIQTVKLAFSVTNNVIRLKPPSSVVSP     832
P troglodytes      IMGIGVDTIDMVMAAQPRRHPSDGAEGQGQGQLLIKTVKLAFSVTNNVIRLKPPSNVVSP     839
B taurus           IMGIGVDTIDMVMAAQPR---SDGAEGQGQGQLLIKTVKLAFSVTNNVIRLKPPSNVVSP     836
                   ***************:*      .. *.********:******************.**** 
H sapiens          LEQALSQHGAHGNNLIAVLAKYIYHKHDPALPRLAIQLLKRLATVAPMSVYACLGNDAAA     896
M musculus         LEQALTQHGAHGNNLIAVLAKYIYHRHDPALPRLAIQLLKRLATVAPMSVYACLGSDAAA     896
R norvegicus       LEQALTQHGAHGNNLIAVLAKYIYHRHDPALPRLAIQLLKRLATVAPMSVYACLGSDAAA     896
G gallus           LEQALTQHGAHGNNLIAVLAKYIYHKHDPALPRLAIQLLKRLATVAPMSVYACLGSDAAA     892
P troglodytes      LEQALSQHGAHGNNLIAVLAKYIYHKHDPALPRLAIQLLKRLATVAPMSVYACLGNDAAA     899
B taurus           LEQALTQHGAHGNNLIAVLAKYIYHKHDPALPRLAIQLLKRLATVAPMSVYACLGSDAAA     896
                   *****:******************:*****************************.**** 
H sapiens          IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSDGSKEFSLGMW     956
M musculus         IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSNGSKEFSLGVW     956
R norvegicus       IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSSGSKEFSLGVW     956
G gallus           IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSDGSKEFSLGEW     952
P troglodytes      IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSDGSKEFSLGMW     959
B taurus           IRDAFLTRLQSKIEDMRIKVMILEFLTVAVETQPGLIELFLNLEVKDGSDGSKEFSLGVW     956
                   ************************************************.******** * 
H sapiens          SCLHAVLELIDSQQQDRYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPKFWENLTSP    1016
M musculus         SCLHVVLELIDSQQQDRYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPKFWENLTSP    1016
R norvegicus       SCLHVVLELIDSQQQDRYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPRFWENLTSP    1016
G gallus           SCLQVVLELIDSKQQERYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPKFWENLTNP    1012
P troglodytes      SCLHAVLELIDSQQQDRYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPKFWENLTSP    1019
B taurus           SCLHAVLELIGSQQQDRYWCPPLLHRAAIAFLHALWQDRRDSAMLVLRTKPKFWENLTSP    1016
                   ***:.******.*:**:************************************:******.* 
H sapiens          LFGTLSPPSETSEPSILETCALIMKIICLEIYYVVKGSLDQSLKDTLKKFSIEKRFAYWS    1076
M musculus         LFGTLSPPSETSEPSVLETCALIMKIICLEIYYVVKGSLDQSLKDTLKKFSSEKRFAYWS    1076
R norvegicus       LFGTLSPPSETSEPSILETCALIMKIICLEIYYVVKGSLDQSLKDTLKKFSSEKRFAYWS    1076
G gallus           LFGTLPPPSETSELSVLDTCALIMKIICLEIYYVVKGSLDQSLKDTLKNFSTKKRFAYWS    1072
P troglodytes      LFGTLSPPSETSEPSILETCALIMKIICLEIYYVVKGSLDQSLKDTLKKFSIEKRFAYWS    1079
B taurus           LFGTLSSPSETSEPSILETCALIMKIICLEIYYVVKGSLDQSLKDTLKKFSSEKRFAYWS    1076
                   *****  ****** *:*:*******************************:** :******* 
H sapiens          GYVKSLAVHVAETEGSSCTSLLEYQMLVSAWRMLLIIATTHADIMHLTDSVVRRQLFLDV    1136
M musculus         GYVKSLAVYMADTEGSSCTSLLEYQMLVSAWRILLIIAASHADVMHLTDMAVRRQLFLDV    1136
R norvegicus       GYVKSLAVYMAETEGNSCTSLLEYQMLVSAWRTLLIVATSHADIMNLTDVAVRRQLFLDV    1136
G gallus           EYVKSLAYYVAETEGSSCASMTEYQMLISAWRMLLIISTSHADIMHLTDPAVHQQLFLDV    1132
P troglodytes      GYVKSLAVHVAETEGSSCTSLLEYQMLVSAWRMLLIIATTHADIMHLTDSVVRRQLFLDV    1139
B taurus           GYVKSLAVHMADTEGSSCTSLVEYQMLVSAWRMFLIIATSHAEITHLTDSAVRRQLFLDV    1136
                    ****** ::*:***.**:*: ***** :**** :**:::::**:: :*** .*::****** 
H sapiens          LDGTKALLLVPASVNCLRLGSMKCTLLLILLRQWKRELGSVDEILGPLTEILEGVLQADQ    1196
M musculus         LDGTKALLLVAASVNCLRLGSMMCTLLLILLRQWKRELGAVEKILGPLTEILEGVLQADQ    1196
R norvegicus       LDGTKALLLVATSVNCLRLGSMMCTLLLILLRQWKRELGAVDKILGPLTEVLEGVLQADQ    1196
G gallus           LNGTKALLLVPMSVSCLQLGSMLCTLLLILLKQWKSELGSVDKIINSLTQILEGVLQADQ    1192
P troglodytes      LDGTKALLLVPASVNCLRLGSMKCTLLLILLRQWKRELGSVDEILGPLTEILEGVLQADQ    1199
B taurus           LEGTKALLLVPTSVNCLRLGSMTCTLLLILLRQWKRELGSVDDILGPLTEILEGVLQADQ    1196
                   *:********  **.**:**** ********:*** ***:*:.*:. **::********* 
```

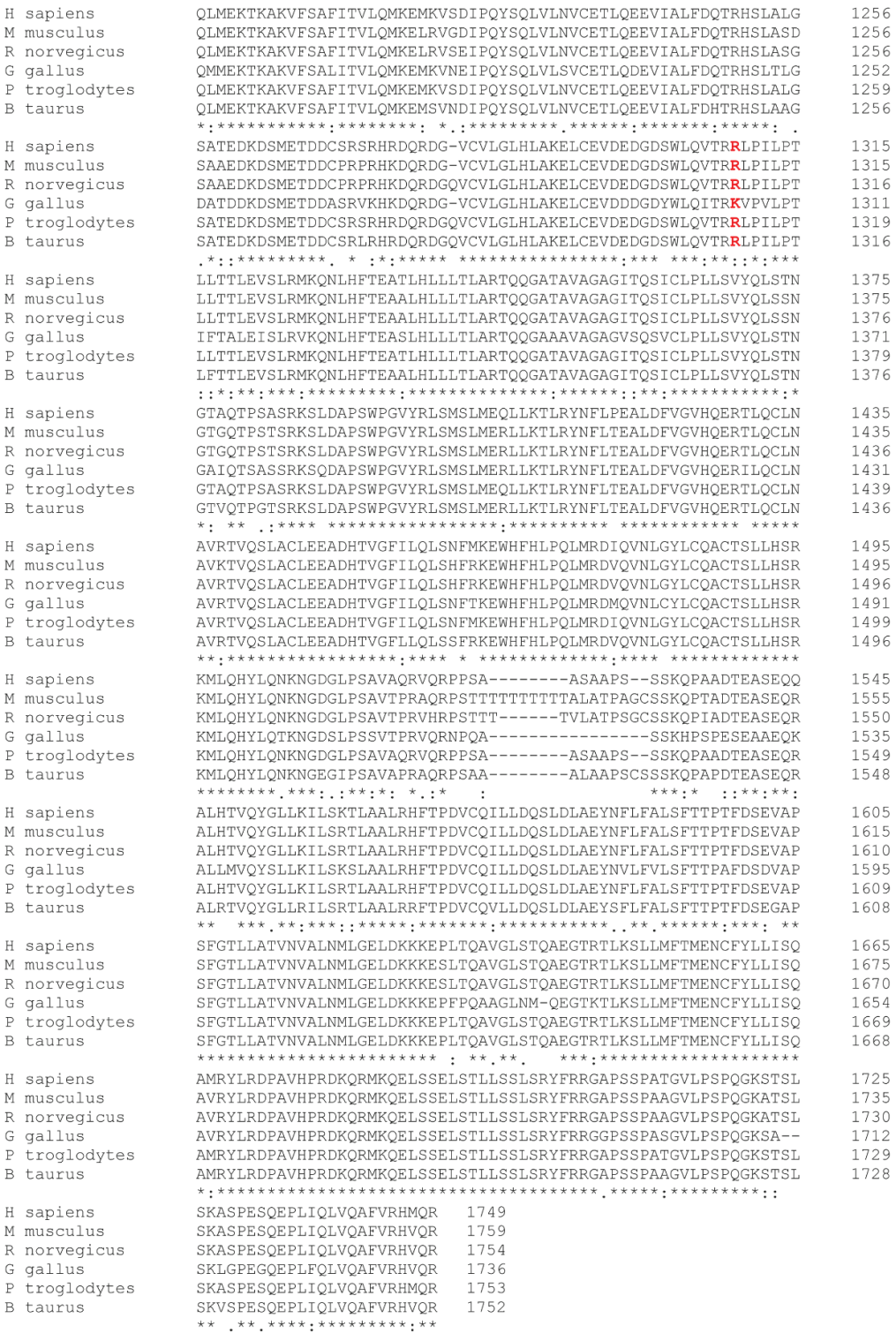


Figure H4: **Multiple sequence alignment showing evolutionary conservation of NUP188.**

Clustal Omega (URL: https://www.ebi.ac.uk/Tools/msa/clustalo/) was used to align the NUP188 protein sequences using the RefSeqs of *H sapiens* (NP_056169.1), *M musculus* (NP_938046.2), *R norvegicus* (XP_006233967.1), *G gallus* (NP_001025861.1), P troglodytes (XP_016801867.2), *B taurus* (NP_001095308.1).

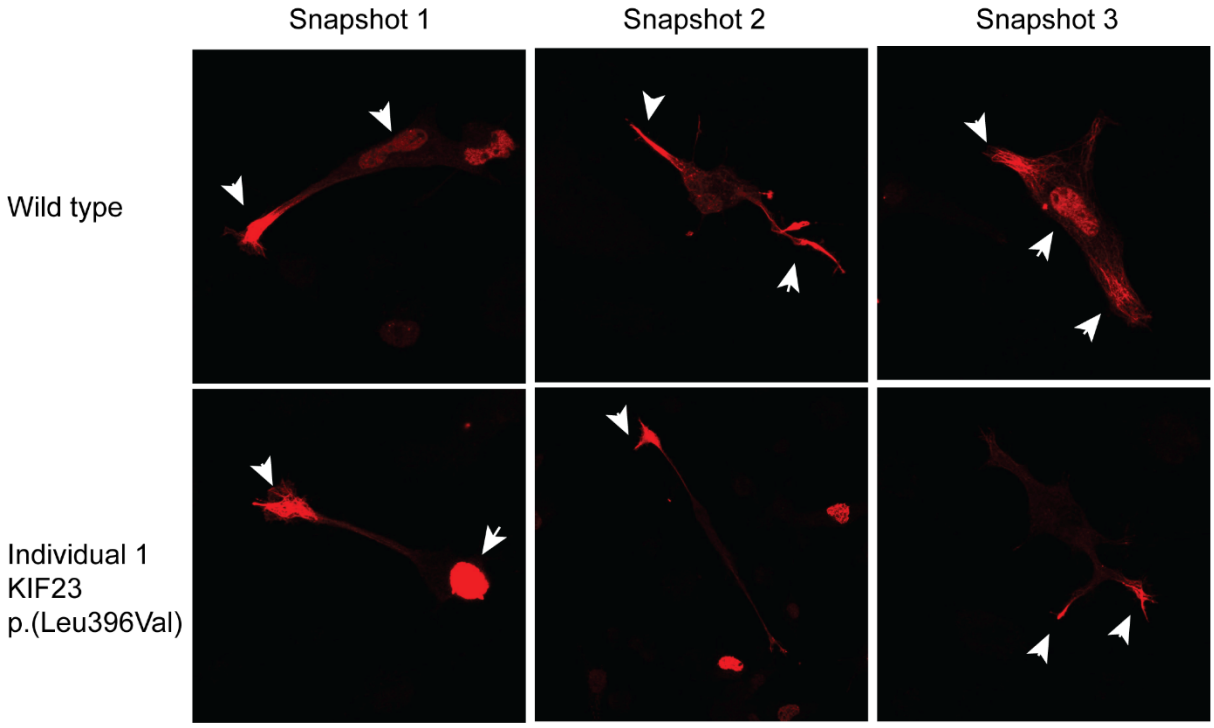


Figure H5: **Neurite tip accumulation assay for *MKLP1*/*KIF23* variant in individual 1**.

The representative images showing the localization of human pmCherry-C1-MKLP1 motors (red) in Neuro2a cells indicated using white arrows. Although, a proportion of expressed protein is localized within the nucleus, the motor for both the wild type and variant construct is able to translocate towards the neurite tip within 24 hours. Qualitative studies did not show any difference in the pattern of localization between control and individual 1's fibroblasts.

Appendix I

Conference proceedings

Conference oral presentations

Kaur, S., Van Bergen, N. J., Gold, W. A., Verhey, K., Boyle, L., Nowell, C. J., . . . Christodoulou, J. (2019). *Emerging role of "kinesinopathies" in genetically undiagnosed Rett Syndrome.* Poster presentation at the MDHS ECR Network Symposium, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., & Christodoulou, J. (2019). *Traffic Jam in brain cells.* Paper presented at the 3 Minute Thesis, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Verhey, K., Boyle, L., Howden, S., . . . Christodoulou, J. (2018). *Identifying novel disease genes in genetically undiagnosed probands with Rett syndrome*. Paper presented at the Human Genetics Society of Australasia, Sydney, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Verhey, K., Boyle, L., Chung, W. K., . . . Christodoulou, J. (2019). *Kinesin family member 1A: A new disease gene for Rett syndrome?* Paper presented at the Annual student research symposium, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Eggers, S., Lunke, S., & Christodoulou, J. (2017). *Identifying novel disease genes in genetically undiagnosed patients with Rett/ Rett-like syndrome.* Paper presented at the The Australian Society for Medical Research, Melbourne, Australia.

Kaur, S., Schonewolf-Greulich, B., Van Bergen, N. J., Gold, W. A., Miller, D., Roscioli, T., . . . Christodoulou, J. (2017). *Uncovering primary defects in kinesins in Rett syndrome and related disorders.* Paper presented at the Annual student research symposium, Melbourne, Australia.

Conference poster presentations

Kaur, S., Van Bergen, N. J., Gold, W. A., Verhey, K., Boyle, L., Howden, S., . . . Christodoulou, J. (2019). *Kinesinopathies: Emerging role in genetically undiagnosed Rett Syndrome.* Paper presented at the American Society of Human Genetics, Houston, USA.

Kaur, S., Van Bergen, N. J., Gold, W. A., Howden, S., Dottori, M., Kupier, M., . . . Christodoulou, J. (2018c). *Uncovering novel disease genes in Rett syndrome patients awaiting genetic diagnosis for years.* Paper presented at the Annual student research symposium, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Howden, S., Dottori, M., Kupier, M., . . . Christodoulou, J. (2018b). *Kinesin deficiency in Rett-Syndrome like patients: 3D-cellular modelling and in vitro functional studies.* Paper presented at the International Society for Stem Cell Research, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Howden, S., Dottori, M., Kupier, M., . . . Christodoulou, J. (2018a). *Kinesin deficiency in Rett-syndrome like patients: 3D-cellular modelling and in vitro functional studies.* Paper presented at the Rare disease Australian functional genomics conference, Melbourne, Australia.

Kaur, S., Van Bergen, N. J., Gold, W. A., Howden, S., Dottori, M., Kupier, M., . . . Christodoulou, J. (2017). *Identifying novel disease genes in genetically undiagnosed patients with Rett/ Rett-like syndrome.* Paper presented at the Murdoch Children Research Institute International review meeting, Melbourne, Australia.

Kaur, S., Schonewolf-Greulich, B., Van Bergen, N. J., Gold, W. A., Miller, D., Roscioli, T., . . . Christodoulou, J. (2017). *Kinesinopathies: a new cause of Rett syndrome?* Paper presented at the BiomedLINK, Melbourne, Australia.

Kaur, S., Schonewolf-Greulich, B., Van Bergen, N. J., Gold, W. A., & Christodoulou, J. (2016). *Identifying novel disease genes in genetically undiagnosed patients with Rett/ Rett-like syndrome.* Paper presented at the Annual student research symposium, Melbourne, Australia.