Dissecting the genetic basis of neurodevelopmental disorders and demyelinating neuropathies

by Stephanie Efthymiou

Supervised by Professor Henry Houlden & Dr. Conceição Bettencourt

A thesis submitted to University College London for the degree of Doctor of Philosophy

Department of Neuromuscular disorders, UCL Queen Square Institute of Neurology, University College London

Declaration

I, Stephanie Efthymiou, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Some chapters are a slightly modified version of the work published by myself, and has been reproduced here with the permission of the copyright holder. Collaborative work is also indicated in this thesis.

Abstract

The understanding of the pathophysiology of most rare, complex neurological disorders has been elusive, especially in the case of complex demyelinating neuropathies and neurodevelopmental disorders. In my work, I learnt to employ two main techniques that will help advance the search for better understanding of neurodevelopmental disorders: next generation sequencing and functional validation of rare genetic variants.

The main aim of my research was to establish the genetic diagnosis in several patients affected by complex syndromes such as peripheral neuropathy with central nervous system involvement (Chapter 3), neurodevelopmental disorders (Chapter 4) and epilepsy (Chapter 5). The phenotypic and genotypic correlations of identified gene variants were investigated in these chapters and is a profound theme in my project. To achieve this, an integrated approach combining next generation sequencing (NGS) technology, homozygosity mapping, array genotyping, traditional Sanger sequencing and functional experiments was undertaken.

Firstly, I describe the work performed in an attempt to identify the causative gene in a cohort of young children presented with an early-onset hereditary form of chronic inflammatory demyelinating polyneuropathy with a central and peripheral involvement. My key findings were that: i) neurofascin is the first gene causally responsible for an inherited disorder that resembles CIDP, ii) this is the largest clinical cohort to date of patients with NFASC mutations with 10 individuals, and iii) the functional evidence implicate the major protein isoforms, which were also shown to be the main targets for the autoantibodies in CIDP pathogenesis.

Secondly, I describe the work done on various neurodevelopmental disorder (NDD) genes, with particular focus on a newly identified gene presenting with a complex neurodevelopmental phenotype comprised of developmental delay, epilepsy, and/or a demyelinating neuropathy. My key findings were that: i) *NARS1*, a cytoplasmic aminoacyl-tRNA synthetase enzyme can be causative for this disorder by either a *de-novo* heterozygous or a biallelic inheritance mode, ii) functional investigations showed reduced aminoacylation activity in the disease-associated biallelic mutations using fibroblasts and iNPCs transcriptomics, suggesting that the majority of *NARS1* mutations cause a loss of function of the protein by reduced expression and disruption of dimer formation suggesting a loss-of-function mechanism, and iii) increased yeast growth in the disease-associated heterozygous mutations showing near normal protein expression are suggestive of a gain-of-function mechanism.

Finally, I describe the work done on two relative new genes (*PIGS* and *TARS1*) in an attempt to expand the patient phenotypic spectrum, as well as an interesting candidate gene (*SLITRK3*) linked with epilepsy. I present my understanding for disease-gene discovery that will enable me and other members of the neurogenetics field to identify disease-mechanisms and address important gaps of translational research into rare neurological diseases such as those described in this thesis.

Table of Contents

Impact statement

The impact of the work presented in this thesis is linked to novel disease discovery and its benefits towards patients and public health as well as potential therapeutic advancements and the global academic and clinical communities.

NFASC-associated neurodevelopmental impairment and demyelination is now considered a new member of the 'hereditary nodo-paranodopathies' and during the time we published the NFASC-related disease[1] together with the recent case reports [2-4], the Online Mendelian Inheritance in Man (OMIM) catalogued it as a new disease - Neurodevelopmental disorder with central and peripheral motor dysfunction (OMIM: 618356). Similarly, the discovery and functional validation of the *NARS1* gene has brought the number of characterised aminoacyl-tRNA-synthetases from 35 to 37 showing the importance of these enzymes in human neurodevelopment. The vital functional validation in our paper[5] and my participation in various national and international conferences raising awareness about these new diseases has led clinicians around the world to include *NFASC* and *NARS1* in their diagnostic analysis. Consequently, around 5 more *NFASC* and 7 more *NARS1* cases have been identified worldwide with follow-up studies being underway.

Remarkably, the identification of these genes has paved the way towards the development of CRISPR/Cas9 heterozygous knock-in and homozygous knock-out animal models. This is the next important step in understanding the molecular rationale of these genetic disease-causing variants. Considering the high number of individuals and variants identified in this thesis, the addition of these important genes

to genetic testing panels for children and young adults presenting with NDD, epilepsy, and/or a demyelinating neuropathy will prove of clinical benefit.

The recent disease discovery work presented in this thesis as part of the SYNAPS collaborative project has placed UCL and Prof. Houlden's lab at the epicentre of collaborative research. The work generated as part of this thesis but also the many collaborative projects with departments within UCL, other institutions in the UK as well as globally with neurologists, bioinformaticians and basic scientists, have generated a model for future research into complex rare disorders. This includes a vast amount of sequencing data run and maintained by the Queen Square Genomics group, a SYNAPS registry of around 20,000 samples with patient-derived cell models, providing the ideal infrastructure to attract collaborations from pharmaceutical industries aiming to work on the discovery and validation of targeted therapies.

The research described in this thesis has been presented orally at the 2019 Peripheral Nerve Society meeting, the European Human Genetics Virtual Conference ESHG 2020 and has been awarded three travel grants (Genetics Society for £585 in 2019, Guarantors of Brain for £600 in 2017 and £750 in 2019). In addition, the *NFASC* work has been awarded a seeding grant from the SOLVE-RD – Solving the Unsolved Rare Diseases" project, to generate and characterize a genetically altered *NFASC* mouse model. The *NFASC* paper has also been described in a scientific commentary in Brain[6]. These honours together with the publications which I have co-authored, highlight the commitment I have for research the work I have contributed in the clinical fields of neurology and neurogenetics.

Acknowledgements

I am extremely grateful to my supervisor Professor Henry Houlden for accepting me in his lab back in May 2015, and for providing me with the opportunity to be an important member of the neurogenetics lab. Prof. Houlden has allowed me to pursue my research ideas through his funding while always providing continuous mentorship, and guidance. I am also grateful to my secondary supervisor Dr. Conceição Bettencourt for teaching me vital lab techniques, encouraging me and providing additional support and ideas but also for supporting my career development.

I would like to thank the following funding bodies for providing the financial support behind my projects and facilitating research exchange and travel to conferences: Wellcome Trust, Horizon 2020, Medical Research Council (MRC), Guarantors of Brain but also the A. G. Leventis Foundation for covering my PhD fees for 2 years. Thanks to new and old members of the lab, to Dr. Vincenzo Salpietro, Dr. Sarah Wiethoff, Dr. Nancy Malintan, Dr. Emer O'Connor, Dr. Reza Maroofian, Roisin Sullivan, Dr. Thomas Bourinaris, Dr. Viorica Chelban, Dr. Rauan Kairzhanov, Natalia Dominik, Hau Ying Yip (Janice), Dr. Andrea Cortese, Dr. Menelaos Pipis, Marilena Christoforou, Ben O'Callaghan, Dr. Mina Ryten, Ms. Deborah Hughes, Dr. Lucia Schottlaender, Dr. Niccolo Mencacci, Dr. David Lynch, Dr. Enrico Bugiardini, Dr. Andreea Manole and Dr. James Polke for their help and guidance at IoN. You have all helped me a lot during the last years, not just as colleagues but also as friends.

Special thanks to Hallgeir Jonvick, David Murphy, Dr. Jana Vandrovcova and Dr. Alan Pittman for helping with data transfer and NGS analysis strategies at the IoN,

but also Mark Gaskin, Alkyoni Athanasiou Fragkouli and Eloise Tribollet for their continuous willingness to help with sample requests and fibroblast cultures.

Thanks also to all the international collaborators who contributed to many of the projects contained in this thesis, including the Morocco group (Prof Yamna Kriouile, Mhammed Aguennouz and Dr Mohamed El Khorassani), Montpelier (Dr. Jerome Devaux), Naples (Prof. Maria Nolano), Trieste (Dr. Sara Fortuna, Dr. Rita de Zorzi), Kings College London (Prof. Corinne Houart, Dr. Patricia M Gordon), Pakistan (Dr. Tipu Sultan, Dr. Faisal Zafar, Dr. Shazia Maqbool), Jordan (Dr. Issam Khawaja), Amsterdam (Dr. Marisa Mendes), Dusseldorf (Dr. Felix Distelmaier), Cambridge (Dr. Rita Horvath), San Diego, CA (Prof. Joseph G. Gleeson), Strasbourg (Dr. Hubert Becker and Dr. Jean-Louis Mandel), Nijmegen (Dr. David A. Koolen), Bordeaux (Dr. Marie Sissler), GeneDx (Dr. Erin Torti), NIH (Dr. Wenyan Han, Dr. Lu Wei, Dr. Jun Li) and all the members of the SYNaPS Study Group.

On a personal note, I want to thank my family, my mother, father and brother in Cyprus who have always been here for me and have provided uncountable moral support. This thesis is dedicated to my two parents who always encouraged me to follow my dreams and ambitions; my loving mother and to my late father, who left us far too early.

Finally, thanks to my partner Dio for patiently standing by my side with constant love and support since the very beginning of this journey. You have always supported without complaining my passion and dedication for this research, which certainly drained a lot of my energy in the past 3 years.

Abbreviations

Abbreviations	Full Description
AA	Amino Acid
AARS1	Alanyl-tRNA Synthetase 1
AARS2	Alanyl-tRNA Synthetase 2
ABI	Applied Biosystems
AD	Autosomal Dominant
ADAR1	Adenosine Deaminase 1
ADCK3	Aarf Domain-Containing Protein Kinase 3
ADCY5	Adenylate Cyclase 5
ADCY6	Adenylate Cyclase 6
ADHD	Attention Deficit/Hyperactivity Disorder
ADP	Adenosine Diphosphate Ribose
ADPRHL2	[Protein ADP-Ribosylarginine] Hydrolase-Like Protein 2
AFO	Ankle Foot Orthosis
AIMP1	Aminoacyl TRNA Synthetase Complex Interacting Multifunctional Protein 1
AIS	Axon Initial Segment
ALPPL2	Alkaline Phosphatase, Placental-Like 2
ALS	Amyotrophic Lateral Sclerosis
Amp	Amplitude
AMPA	A-Amino-3-Hydroxy-5-Methyl-4-Isoxazolepropionic Acid
AMPAR	A-Amino-3-Hydroxy-5-Methyl-4-Isoxazolepropionic Acid Receptor
ANNOVAR	Annotation Of Variants
ANOVA	Analysis Of Variance
APGAR	Appearance, Pulse, Grimace, Activity, And Respiration
APOE	Apolipoprotein E
AR	Autosomal Recessive
ASD	Autism Spectrum Disorder
ASXL1	ADDITIONAL SEX COMBS-LIKE 1
ATF	Aminoterminal Fragment
ATF4	Activating Transcription Factor 4
ATP	Adenosine Triphosphate
BBS2	Bardet-Biedl Syndrome 2
BCA	Bicinchoninic Acid
BCIP	5-Bromo-4-Chloro-3-Indolyl Phosphate
BNPAGE	Blue Native Polyacrylamide Gel Electrophoresis
BOS	Bohring-Opitz Syndrome
bp	Basepair
BRAINEAC	Brain Eqtl Almanac
BRPF1	Bromodomain And PHD Finger Containing 1
BWA	Burrows-Wheeler Aligner
CA1	Cornu Ammonis 1
CACNA1A	Calcium Voltage-Gated Channel Subunit Alpha1 A

CADASIL	Cerebral Autosomal Dominant Arteriopathy With Subcortical Infarcts And Leukoencephalopathy
CADD	Combined Annotation Dependent Depletion
Nrcam	Neuronal Cell Adhesion Molecule
M-CAP	Mendelian Clinically Applicable Pathogenicity Score
CARS1	Cysteinyl-TRNA Synthetase 1
CARS2	Cysteinyl-TRNA Synthetase 2
CASPR1	Alias for Contactin Associated Protein 1
CASPR2	Alias for Contactin Associated Protein 2
CCDS	Consensus Coding Sequence
CDS	Coding Sequence
CEX	Coding Exome Oligos
CGH	Comparative Genomic Hybridization
CHAPS	3-((3-Cholamidopropyl) Dimethylammonio)-1-Propanesulfonate
CIDP	Chronic Inflammatory Demyelinating Polyneuropathy
CLN5	Ceroid-Lipofuscinosis Neuronal Protein 5
CMT	Charcot Marie Tooth
CNS	Central Nervou System
CNTN1	Contactin 1
CNTN2	Contactin 2
CNTNAP	Contactin Associated Protein 1
CNTNAP1	Contactin Associated Protein 2
CNV	Copy Number Variance
COLIV	Collagen IV
CPDB	Consensuspathdb
CRBL	Cerebellar Cortex
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSF	Cerebrospinal Fluid
CSV	Comma-Separated Values
CTRL	Control
DABCO	1,4-Diazabicyclo[2.2. 2]Octane
DAPI	4′,6-Diamidino-2-Phenylindole
DARS1	Aspartyl-Trna Synthetase 1
DARS2	Aspartyl-Trna Synthetase 2
DD	Developmental Delay
DDX	DEAD-Box Helicase
DEAD	Asp-Glu-Ala-Asp
DIG	Digoxigenin
DIV	Days In Vitro
DMD	Duchene Muscular Dystrophy
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNMT3A	DNA Methyltransferase 3 Alpha
DNQX	6,7-Dinitroquinoxaline-2,3-Dione

DPX	Distyrene, Plasticizer, Xylene
DTR	Deep Tendon Reflex
DTT	Dithiothreitol
EA	Episodic Ataxia
EARS2	Glutamyl-TRNA Synthetase 2
EBV	Epstein - Barr Virus
ECACC	European Collection Of Authenticated Cell Cultures
ECG	Electrocardiogram
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic Acid
EEG	Electroencephalogram
EGTA	Egtazic Acid
EIEE	Early Infantile Epileptic Encephalopathy
ELAV	Mbryonic Lethal Abnormal Visual System
EMG	Electromyography
EMM	Edinburgh's Minimal Medium |S.Pombe Growth Medium
EPRS1	Glutamyl-Prolyl-Trna Synthetase 1
ERP3	End Repair Mix 3
ESP6500	Exome Sequencing Project V. 6500
FARS2	Phenylalanyl-Trna Synthetase 2
FARSA	Henylalanyl-TRNA Synthetase Subunit Alpha
FARSB	Henylalanyl-TRNA Synthetase Subunit Beta
FATHMM	Functional Analysis Through Hidden Markov Models
FBS	Foetal Bovine Serum
FCTX	Frontal Cortex
FDR	False Discovery Rate
FITC	Fluorescein Isothiocyanate
FL	Fluorescence
FLAIR	Fluid-Attenuated Inversion Recovery
FMRP	Fragile X Mental Retardation Protein
FOSB	FBJ Murine Osteosarcoma Viral Oncogene Homolog B
FPKM	Fragments Per Kilobase Of Transcript Per Million Mapped Reads
FRAXA	Fragile X
FS	Frameshift
FTD	Frontotemporal Dementia
FTLD	Frontotemporal Lobar Degeneration
FUS	Fused in Sarcoma/Translocated in Sarcoma
GABA	Gamma-Aminobutyric Acid
GAD1	Glutamic Acid Decarboxylase 1
GAIT	Gamma Interferon Inhibitor of Translation
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GARS1	Glycyl-TRNA Synthetase 1
GATK	Genome Analysis Tool Kit
GCN2	General Control Nonderepressible 2

GDAP2	Ganglioside Induced Differentiation Associated Protein 2
GDD	Global Developmental Delay
GENSAT	Gene Expression Nervous System Atlas
GERP	Enomic Evolutionary Rate Profiling
GFP	Green Fluorescent Protein
GJB1	Gap Junction Beta-1 Protein
GLDN	Gliomedin
GLUT1	Glucose Transporter 1
GNB5	G Protein Subunit Beta 5
GO	Gene Ontology
GPI	Glycosyl Phosphatidyl Inositol
Grch37	Genome Reference Consortium Human Genome Build 37
GRIA2	Glutamate Ionotropic Receptor AMPA Type Subunit 2
GRIN1	Glutamate Ionotropic Receptor NMDA Type Subunit 1
GST	Glutathione S-Transferase
GTC	Generalised Tonic Clonic
Gtex	Genotype Tissue Expression Consortium
GTP	Guanosine Triphosphate
HARS1	Histidyl-Trna Synthetase 1
HARS2	Histidyl-Trna Synthetase 2
HEK	Human Embryonic Kidney
HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic Acid
HGVS	Human Genome Variation Society
HIPP	Hippocampus
HMN	Hereditary Motor Neuropathy
HOX	Homeobox
HRP	Horseradish Peroxidase
HSN	Hereditary Sensory Neuropathy
IARS1	Isoleucyl-TRNA Synthetase 1
IARS2	Isoleucyl-TRNA Synthetase 2
ID	Intelectual Disability
IMPC	International Mouse Phenotyping Consortium
INDEL	Insertion/Deletion
IPSC	Induced Pluripotent Stem Cell
IRB	Institutional Review Board
ISCN	International System For Human Cytogenetic Nomenclature
ISR	Integrated Stress Response
IUGR	Intrauterine Growth Restriction
JAM2	Junctional Adhesion Molecule 2
KARS1	Lysyl-TRNA Synthetase 1
KAT6A	Lysine Acetyltransferase 6A
KCNK4	Potassium Channel Subfamily K Member 4
KCNQ2	Potassium Voltage-Gated Channel Subfamily Q Member 2
KIF1A	Kinesin Family Member 1A

LARS1	Leucyl-TRNA Synthetase 1
LARS2	Leucyl-TRNA Synthetase 2
LCL	Lymphoblastoid Cell Lines
LINCS	Linear Constraint Solver
LRR	Leucine-Rich Repeat
LRT	Likelihood Ratio Test
MAF	Minor Allele Frequency
MARS1	Methionyl-TRNA Synthetase 1
MARS2	Methionyl-TRNA Synthetase 2
MBP	Myelin Binding Protein
MEDU	Medulla
MEM	Modified Eagle Medium
MIM	Mendelian Inheritance In Man
MLPA	Multiplex Ligation-Dependent Probe Amplification
MM	Master Mix
MND	Motor Neurone Disease
MRC	Medical Research Council
MRI	Magnetic Resonance Imaging
MS	Milliseconds
MTMR2	Myotubularin Related Protein 2
MYORG	Myogenesis Regulating Glycosidase
NARS1	Asparaginyl-TRNA Synthetase 1
NARS2	Asparaginyl-TRNA Synthetase 2
NBIA	Neurodegeneration With Brain Iron Accumulation
NBT	Nitro Blue Tetrazolium
NCBI	National Center For Biotechnology Information
NCS	Nerve Conduction Studies
NCV	Nerve Conduction Velocity
NDD	Neurodevelopment
NEB	New England Biolabs
NFASC	Neurofascin
NFASC155	Neurofascin 155
NFASC186	Neurofascin 186
NFASC3	Neurofascin Isoform 3
NGS	Next-Generation Sequencing
NKX6	NK6 Homeobox 2
NL2	Neuroligin2
NMD	Neuromuscular Disorders
NOPS	NONA/Paraspeckle
NOTCH3	Notch Receptor 3
NP40	Nonyl Phenoxypolyethoxylethanol
NPC	Neural Progenitor Cells
NPT	Constant number of atoms, volume and temperature (canonical ensemble)
NTNG2	Netrin G2

NVT	Constant number of atoms, volume and temperature (canonical ensemble)
OCTX	Occipital Cortex
OFC	Occipital Frontal Circumference
OFD5	Orofaciodigital Syndrome Type 5
OGDHL	Oxoglutarate Dehydrogenase L
ORC1	Origin Recognition Complex Subunit 1
PAGE	Polyacrylamide Gel Electrophoresis
PANK2	Pantothenate Kinase 2
PAP	Postaxial Polydactyly
PARP1	Poly(ADP-Ribose) Polymerase 1
PARS2	Prolyl-TRNA Synthetase 2
PBL	Peripheral Blood Lymphocyte
PBS	Phosphate Buffered Saline
PBST	PBS/0.1% Tween-20
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDXK	Pyridoxal Kinase
PEHO	Progressive Encephalopathy With (O)Edema, Hypsarrhythmia And Optic Atrophy
PFA	Paraformaldehyde
PGP	P-Glycoprotein
PHD	Plant Homeodomain
PIGK	Phosphatidylinositol Glycan Anchor Biosynthesis Class K
PIGS	Phosphatidylinositol Glycan Anchor Biosynthesis Class S
PLP	Pyridoxal (PL) 5′-Phosphate
PNS	Peripheral Nervous System
PPC	PCR Primer Cocktail
PRRT2	Proline-Rich Transmembrane Protein 2
PRUNE	Prune Exopolyphosphatase
PUTM	Putamen
PVDF	Polyvinylidene Difluoride
PVL	Periventricular Leukomalacia
QARS1	Glutaminyl-TRNA Synthetase 1
QC	Quality Control
RARS1	Arginyl-TRNA Synthetase 1
RARS2	Arginyl-TRNA Synthetase 2
RFC1	Replication Factor C Subunit 1
RMSD	Simply Root-Mean-Square Deviation
RMSF	Root Mean Square Fluctuations
RNA	Ribonucleic Acid
RT	Room Temperature
SARS1	Seryl-TRNA Synthetase 1
SARS2	Seryl-TRNA Synthetase 2
SCA14	Spino-Cerebellar Ataxia 14

SD	Standard Deviation
SFPQ	Splicing Factor Proline And Glutamine Rich
SHH	Sonic Hedhehog
SHIRPA	Smithkline Beecham, Harwell, Imperial College, Royal London Hospital, Phenotype Assessment
SIFT	Sorting Intolerant From Tolerant
SLITRK3	SLIT And NTRK Like Family Member 3
SMA	Spinal Muscular Atrophy
SNARE	Soluble N-Ethylmale-Imide-Sensitive Factor-Attachment Protein Receptors
SNIG	Substantia Nigra
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variant
SOC	Super Optimal Broth With Catabolite Repression
SOX1	SRY-Box Transcription Factor 1
SP6	Spastic Paraplegia-6
SPG11	Spastic Paraplegia-11
SSC	Saline-Sodium Citrate
SSR	Sympathetic Skin Response
STAR	Spliced Transcripts Alignment To A Reference
SVF	Stromal Vascular Fraction
SYNE1	Spectrin Repeat Containing Nuclear Envelope Protein 1
TARS1	Threonyl-Trna Synthetase 1
TARS2	Threonyl-Trna Synthetase 2
TBCD	Tubulin Folding Cofactor D
TBE	Tris/Borate/EDTA
TBST	Tris-Buffered Saline, 0.1% Tween® 20 Detergent
TCTX	Temporal Cortex
TGS	Thrrs, Gtpase, And Spot
TTD	Trichothiodystrophy
TTN	Titin
TTX	Tetrodotoxin
TWIST2	Twist Family BHLH Transcription Factor 2
UGP2	UDP-Glucose Pyrophosphorylase 2
UTR	Untranslated Region
UV	Ultraviolet
VAMP1	Vesicle Associated Membrane Protein 1
VAMP2	Vesicle Associated Membrane Protein 2
VARS1	Valyl-TRNA Synthetase 1
VARS2	Valyl-TRNA Synthetase 2
VCF	Variant Call Format
VEGFR1	Vascular Endothelial Growth Factor Receptor 1
VEP	Variant Effect Predictor
WARS1	Tryptophanyl-TRNA Synthetase 1
WARS2	Tryptophanyl-TRNA Synthetase 2

WES	Whole Exome Sequencing
WGCNA	Weighted Correlation Network Analysis, Also Known As Weighted Gene Co-Expression Network Analysis
WGS	Whole Genome Sequencing
WHMT	White Matter
XRCC1	X-Ray Repair Cross Complementing 1
YARS1	Tyrosyl-TRNA Synthetase 1
YARS2	Tyrosyl-TRNA Synthetase 2

Web Resources

Online Mendelian Inheritance in Man (OMIM): http://www.omim.org/

Primer 3: http://primer3.ut.ee/

Ensembl: http://www.ensembl.org/

ANNOVAR, www.openbioinformatics.org/annovar

National Center for Biotechnology Information (NCBI), http://www.ncbi.nlm.nih.gov

UCSC Genome Browser: https://genome.ucsc.edu/

Picard: http://broadinstitute.github.io/picard/

Human Genome Variation Society: http://www.hgvs.org

FastQC: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

CPDB web tool: http://cpdb.molgen.mpg.de/

Homozygosity Mapper: http://www.homozygositymapper.org/

NHLBI Exome Variant Server EVS: http://evs.gs.washington.edu/EVS/

1000 Genomes project: www.1000genomes.org

cg69 DB: http://www.completegenomics.com/publicdata/69-Genomes

dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP

Exome Aggregation Consortium database: http://exac.broadinstitute.org/

Mutation Taster ver. 2: http://www.mutationtaster.org/

PolyPhen2: http://genetics.bwh.harvard.edu/pph2/

SIFT: http://sift.jcvi.org/

CADD: http://cadd.gs.washington.edu/home

GATK, Best Practices: https://www.broadinstitute.org/gatk/guide/best-practices.php

PYMOL, www.pymol.org

Allen Brain Atlas, http://www.brain-map.org/

Genomes England Panel App https://panelapp.genomicsengland.co.uk/

List of Tables

List of Figures

List of Publications

(Part of the work described in this thesis has been published in the following articles)

First author publications arising from this thesis:

***shared authorship**

1. Manole A***, Efthymiou S***, O'Connor E*, Mendes MI*, Jennings M*, Maroofian R, Davagnanam I, Mankad K, Lopez MR, Salpietro V, Harripaul R. De-novo and Bi-allelic Pathogenic Variants in NARS1 Cause Neurodevelopmental Delay Due to Toxic Gain-of-Function and Partial Loss-of-Function Effects. The American Journal of Human Genetics. 2020 Jul 31.
2. **Efthymiou S,** Kriouile Y, Salpietro V, Hajar R, Ghizlane Z, Mankad K, El Khorassani M, Aguennouz M, Houlden H, Wiethoff S. A rare PANK2 deletion in the first north African patient affected with pantothenate kinase associated neurodegeneration. Journal of the Neurological Sciences. 2020 Mar 15;410.
3. **Efthymiou S***, Gordon PM*, Salpietro V, Fielding T, Borgione E, Scuderi C, Houlden H, Houart C. Human patient SFPQ homozygous mutation is found deleterious for brain and motor development in a zebrafish model. bioRxiv. 2020 Jan 1.
4. **Efthymiou S,** Salpietro V, Pironti E, Bonsignore M, Ferrazzoli V, Di Rosa G, Houlden H. A de-novo truncating mutation in ASXL1 associated with segmental overgrowth. Journal of Genetics. 2019 Dec 1;98(5):108.
5. **Efthymiou S**, Salpietro V, Malintan NT, Poncelet M, Kriouile Y, Fortuna S, De Zorzi R, Payne K, Henderson LB, Cortese A, Maddirevula S, Alhashmi N, Wiethoff S, Ryten M, Botia JA, Provitera V, Schuelke M, Bettencourt

C,Vandrovcova J, Karimiani EG, Maroofian R, Aguennouz M, El Khorassani M, Schmidts M, Alkuraya FS, Edvardson S, Nolano M, Devaux J, Houlden H. Biallelic Neurofascin variants affect paranodal axoglial junctions causing neurodevelopmental impairment and central and peripheral demyelination. Brain. 2019 Sep 9;142(10):2948-64.

6. **Efthymiou S**, Salpietro V, Bettencourt C, Houlden H. Paroxysmal movement disorder and epilepsy caused by a de-novo truncating mutation in KAT6A. Journal of pediatric genetics. 2018 Sep;7(03):114-6.
7. Salpietro V*, **Efthymiou S*,** Manole A, Maurya B, Wiethoff S, Ashokkumar B, Cutrupi MC, Dipasquale V, Manti S, Botia JA, Vandrovcova J, Bettencourt C, Mankad K, Mukherjee A, Mutsuddi M, Houlden H. A loss-of-function homozygous mutation in DDX59 implicates a conserved RNA helicase in nervous system development and function. *Human Mutation*. 2017;1–6.
8. **Efthymiou S**, Manole A, Houlden H. Next generation sequencing in neuromuscular diseases. Current opinion in neurology. 2016 Oct;29(5):527.

Other publications (co)authored during this thesis:

1. De Nittis P, **Efthymiou S**, Sarre A, Guex N, Chrast J, Putoux A, Sultan T, Alvi JR, Rahman ZU, Zafar F. Inhibition of G-protein signaling in cardiac dysfunction of Intellectual Developmental Disorder with Cardiac Arrhythmia (IDDCA) syndrome. JMG. 2020 Sept 4. (in press)
2. Kumar KR, Cortese A, Tomlinson SE, **Efthymiou S**, Ellis M, Zhu D, Stoll M, Dominik N, Tisch S, Tchan M, Wu KH. RFC1 expansions can mimic hereditary sensory neuropathy with cough and Sjögren syndrome. Brain. 2020 Sep 19.

3. Chen Z, Yan Yau W, Jaunmuktane Z, Tucci A, Sivakumar P, Gagliano Taliun SA, Turner C, **Efthymiou S**, Ibáñez K, Sullivan R, Bibi F. Neuronal intranuclear inclusion disease is genetically heterogeneous. Annals of Clinical and Translational Neurology. 2020 Aug 10.
4. Neuray C, Maroofian R, Scala M, Sultan T, Pai GS, Mojarrad M, El Khashab H, deHoll L, Yue W, Alsaif HS, Zanetti MN, Bello O, Person R, Eslahi A, Khazaei Z, Feizabadi MH, **Efthymiou S**. Early-infantile onset epilepsy and developmental delay caused by bi-allelic GAD1 variants. Brain. 2020 May 1.
5. Breza M, Bourinaris T, **Efthymiou S**, Maroofian R, Athanasiou-Fragkouli A, Tzartos J, Velonakis G, Karavasilis E, Angelopoulou G, Kasselimis D, Potagas C. A homozygous GDAP2 loss-of-function variant in a patient with adult-onset cerebellar ataxia. Brain. 2020 May 18.
6. Azad B, **Efthymiou S**, Sultan T, Scala M, Alvi JR, Neuray C, Dominik N, Gul A, Houlden H, SYNaPS Study Group. Novel likely disease-causing CLN5 variants identified in Pakistani patients with neuronal ceroid lipofuscinosis. Journal of the Neurological Sciences. 2020 Apr 7:116826.
7. Nguyen TT, Murakami Y, Mobilio S, Niceta M, Zampino G, Philippe C, Moutton S, Zaki MS, James KN, Musaev D, Mu W. Bi-allelic Variants in the GPI Transamidase Subunit PIGK Cause a Neurodevelopmental Syndrome with Hypotonia, Cerebellar Atrophy, and Epilepsy. The American Journal of Human Genetics. 2020 Mar 26.
8. Chelban V, Carecchio M, Rea G, Bowirrat A, Kirmani S, Magistrelli L, **Efthymiou S**, Schottlaender L, Vandrovcova J, Salpietro V, Salsano E. MYORG-related disease is associated with central pontine calcifications and atypical parkinsonism. Neurology Genetics. 2020 Apr 1;6(2).

9. Ilyas M, **Efthymiou S**, Salpietro V, Noureen N, Zafar F, Rauf S, Mir A, Houlden H. Novel variants underlying autosomal recessive intellectual disability in Pakistani consanguineous families. *BMC Med Genet*. 2020;21(1):59. 2020 Mar 24.
10. Schottlaender LV, Abeti R, Jaunmuktane Z, Macmillan C, Chelban V, O'Callaghan B, McKinley J, Maroofian R, **Efthymiou S**, Athanasiou-Fragkouli A, Forbes R. Bi-allelic JAM2 Variants Lead to Early-Onset Recessive Primary Familial Brain Calcification. The American Journal of Human Genetics. 2020 Mar 5;106(3):412-21.
11. Bibi F, **Efthymiou S**, Bourinaris T, Tariq A, Zafar F, Rana N, Salpietro V, Houlden H, Raja GK, Saeed S, Minhas NM. Rare novel CYP2U1 and ZFYVE26 variants identified in two Pakistani families with spastic paraplegia. Journal of the Neurological Sciences. 2020 Apr 15;411:116669.
12. Nardello R, Fontana A, Mangano GD, **Efthymiou S**, Salpietro V, Houlden H, Mangano S. Age-dependent epileptic encephalopathy associated with an unusual co-occurrence of ZEB2 and SCN1A variants. Epileptic Disorders. 2020 Feb 6;1(1).
13. Cortese A, Tozza S, Yau WY, Rossi S, Beecroft SJ, Jaunmuktane Z, Dyer Z, Ravenscroft G, Lamont PJ, Mossman S, Chancellor A. Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. Brain. 2020 Feb 1;143(2):480-90.
14. Chelban V, Alsagob M, Kloth K, Chirita-Emandi A, Vandrovcova J, Maroofian R, Davagnanam I, Bakhtiari S, AlSayed MD, Rahbeeni Z, AlZaidan H. Genetic and phenotypic characterization of NKX6-2-related spastic ataxia and hypomyelination. European journal of neurology. 2020 Feb;27(2):334-42.

15. Ilyas M, Mir A, **Efthymiou S**, Houlden H. The genetics of intellectual disability: advancing technology and gene editing. F1000Research. 2020 Jan 16;9(22):22.
16. Ilyas M, Salpietro V, **Efthymiou S**, Bourinaris T, Tariq A, Imdad M, Ahmad A, Ahmad H, Houlden H. Identification of common genetic markers of paroxysmal neurological disorders using a network analysis approach. Neurological Sciences. 2019 Dec 6:1-7.
17. Perenthaler E, Nikoncuk A, Yousefi S, Berdowski WM, Alsagob M, Capo I, van der Linde HC, van den Berg P, Jacobs EH, Putar D, Ghazvini M. Loss of UGP2 in brain leads to a severe epileptic encephalopathy, emphasizing that bi-allelic isoform-specific start-loss mutations of essential genes can cause genetic diseases. Acta Neuropathologica. 2019 Dec 9:1-28.
18. Vandervore LV, Schot R, Milanese C, Smits DJ, Kasteleijn E, Fry AE, Pilz DT, Brock S, Börklü-Yücel E, Post M, Bahi-Buisson N. TMX2 Is a Crucial Regulator of Cellular Redox State, and Its Dysfunction Causes Severe Brain Developmental Abnormalities. The American Journal of Human Genetics. 2019 Nov 14.
19. Dias CM, Punetha J, Zheng C, Mazaheri N, Rad A, **Efthymiou S**, Petersen A, Dehghani M, Pehlivan D, Partlow JN, Posey JE. Homozygous Missense Variants in NTNG2, Encoding a Presynaptic Netrin-G2 Adhesion Protein, Lead to a Distinct Neurodevelopmental Disorder. The American Journal of Human Genetics. 2019 Nov 7;105(5):1048-56.
20. Wang H, Kaçar Bayram A, Sprute R, Ozdemir O, Cooper E, Pergande M, **Efthymiou S**, Nedic I, Mazaheri N, Stumpfe K, Azizi Malamiri R. Genotype-Phenotype Correlations in Charcot-Marie-Tooth disease due to MTMR2

mutations and implications in membrane trafficking. Frontiers in neuroscience. 2019;13:974.

21. Chelban V, Alsagob M, Kloth K, Chirita-Emandi A, Vandrovcova J, Maroofian R, Davagnanam I, Bakhtiari S, AlSayed MD, Rahbeeni Z, AlZaidan H, **Efthymiou S** et al. Genetic and phenotypic characterization of NKX6-2-related spastic ataxia and hypomyelination. European journal of neurology. 2019 Sep 11.
22. Khani M, Taheri H, Shamshiri H, Houlden H, **Efthymiou S**, Alavi A, Nafissi S, Elahi E. Continuum of phenotypes in hereditary motor and sensory neuropathy with proximal predominance and Charcot–Marie–Tooth patients with TFG mutation. American Journal of Medical Genetics Part A. 2019 May 20.
23. Chelban V, Wilson MP, Warman Chardon J, Vandrovcova J, Zanetti MN, Zamba-Papanicolaou E, **Efthymiou S**, Pope S, Conte MR, Abis G, Liu YT. PDXK mutations cause polyneuropathy responsive to PLP supplementation. Annals of neurology. 2019 Jun 11.
24. Salpietro V, Dixon CL, Guo H, Bello OD, Heimer G, Burglen L, Valence S, Torti E, Hacke M, Rankin L, **Efthymiou S**, Tariq H, Colin E, Procaccio V, Striano P, Mankad K, Lieb A, Chen S, Pisani L, Bettencourt C, Manniko R, Manole A, Brusco E, Grosso E, Ferrero GB, Armstrong Moron J, Gueden S, Bar-Yosef O, Tzadok M, Monaghan KG, Santiago-Sim T, Person RE, Willaert R, Cho MT, Vandrovcova J, Yoo Y, Chae JH, Quan Y, Wu H, Wang T, Bernier RA, Xia K, Schneider AL, Boysen K, Muir AM, Van Haeringen A, Ruivenkamp C, Nava C, Heron D, Zara F, Minetti C, Skabar A, Fabretto A, Deciphering Developmental Disorders Study, SYNAPS Study Group, Raspall-Chaure M,

Chez M, Tsai Anne, Fassi E, Shinawi M, Costantino J, De Zorzi R, Fortuna S, Keren B, Bonneau D, Choi M, Benzeev B, Mefford HC, Scheffer IE, Clayton-Smith J, Macaya A, Rothman JE, Eichler EEE, Kullmann DM, Houlden H. 2018. AMPA receptor GluA2 subunit defects are a pivotal cause of neurodevelopmental disorders. Nature communications. 2019 Jul 12;10(1):3094.

25. Savasta S, Bassanese F, Buschini C, Foiadelli T, Trabatti C, **Efthymiou S**, Salpietro V, Houlden H, Simoncelli A, Marseglia GL. Biotin-Thiamine Responsive Encephalopathy: Report of an Egyptian Family with a Novel SLC19A3 Mutation and Review of the Literature. Journal of Pediatric Genetics. 2018 Dec 18.
26. Bell S, Rousseau J, Peng H, Aouabed Z, Priam P, Theroux JF, Jefri M, Tanti A, Wu H, Kolobova I, Silviera H, **Efthymiou S** et al. Mutations in ACTL6B Cause Neurodevelopmental Deficits and Epilepsy and Lead to Loss of Dendrites in Human Neurons. The American Journal of Human Genetics. 2019 Apr 25.
27. Cortese A, Simone R, Sullivan R, Vandrovcova J, Tariq H, Yau, W, Humphrey J, Jaunmuktane J, Sivakumar P, Ilyas M, Polke J, Tribbolet E, Tomaselli PJ, Devigili G, Callegari I, Versino M, Salpietro V, **Efthymiou S**, Kaski D, Wood NW, Andreade SN, Buglo E, Rebelo A, Fratta P, Rossor A, Marques WJ, Zuchner S, Reilly MM, Houlden H. Biallelic Expansion of an intronic Repeat in the RFC1 Gene is a common cause of Late-Onset Ataxia. Nature Genetics 2019 Apr;51(4):649.
28. Salpietro V, Malintan NT, Llano I, Spaeth CG, **Efthymiou S**, Striano P, Cutrupi MC, David E, Di Rosa G, Martinez-Azorin F, Hernandez Elena M,

ZaraF, Minetti C, Deciphering Developmental Disorders Study, SYNAPS Study Group, Bello OD, Mankad K, Macaya A, Prada C, Kullmann DM, Rothman JE, Krishnakumar SS, Houlden H. 2018. Mutations in the neuronal vesicular SNARE VAMP2 affect synaptic membrane fusion and impair human neurodevelopment. Am J Hum Genet. 2019 Mar 28.

29. Pironti E, Granata F, Cucinotta F, Gagliano A, **Efthymiou S**, Houlden H, Salpietro V, Di Rosa G. Electroclinical history of a five-year-old girl with GRIN1-related early-onset epileptic encephalopathy: a video-case study. Epileptic Disorders. 2018 Oct;20(5):423-7.
30. Niccolini F, Mencacci NE, Yousaf T, Rabiner EA, Salpietro V, Pagano G, Balint B, **Efthymiou S**, Houlden H, Gunn RN, Wood N. PDE10A and ADCY5 mutations linked to molecular and microstructural basal ganglia pathology. Movement Disorders. 2018 Oct 21.
31. Ghosh SG, Becker K, Huang H, Salazar TD, Chai G, Salpietro V, Al-Gazali L, Waisfisz Q, Wang H, Vaux KK, Stanley V. Biallelic mutations in ADPRHL2, encoding ADP-ribosylhydrolase 3, lead to a degenerative pediatric stress-induced epileptic ataxia syndrome. The American Journal of Human Genetics. 2018 Sep 6;103(3):431-9.
32. O'Connor E, Vandrovcova J, Bugiardini E, Chelban V, Manole A, Davagnanam I, Wiethoff S, Pittman A, Lynch DS, **Efthymiou S**, Marino S. Mutations in XRCC1 cause cerebellar ataxia and peripheral neuropathy. J Neurol Neurosurg Psychiatry. 2018 Nov 1;89(11):1230-2.
33. Pironti E, Salpietro V, Cucinotta F, Granata F, Mormina E, **Efthymiou S**, Scuderi C, Gagliano A, Houlden H, Di Rosa G. A novel SLC1A4 homozygous mutation causing congenital microcephaly, epileptic encephalopathy and

spastic tetraparesis: a video-EEG and tractography–case study. Journal of neurogenetics. 2018 Jul 5:1-6.

34. Salpietro V, Manole A, **Efthymiou S**, Houlden H. *A review of copy number variants in inherited neuropathies. Curr Genomics. 2017*
35. Chelban V, Wiethoff S, Fabian-Jessing BK, Haridy NA, Khan A, **Efthymiou S**, Becker EB, O'Connor E, Hersheson J, Newland K, Hojland AT. Genotype-phenotype correlations, dystonia and disease progression in spinocerebellar ataxia type 14. Movement Disorders. 2018 Jul;33(7):1119-29.
36. Salpietro V, Perez-Dueñas B, Nakashima K, San Antonio MV, Manole A, **Efthymiou S,** Bettencourt C, Mencacci NE, Klein C, Kelly MP, Davies CH, Kimura H, Macaya A, Houlden H. A homozygous loss-of-function mutation in PDE2A associated to infantile-onset hereditary chorea. Mov Dis 2018 Mar;33(3):482-488
37. Bettencourt C*, Salpietro V*, **Efthymiou S**, Chelban V, Hughes D, Pittman AM, Federoff M, Bourinaris T, Spilioti M, Deretzi G, Kalantzakou T, Houlden H, Singleton AB, Xiromerisiou G. Genotype-Phenotype correlations and expansion of the molecular spectrum of *AP4M1*-related Hereditary Spastic Paraplegia. Orphanet J Rare Dis 2017.

Chapter 1. General introduction

1.1 Next generation sequencing techniques in neuromuscular diseases

Neuromuscular disorders (NMD) are a broad group of conditions that affect muscles (myopathies, dystrophies, ion channel diseases and malignant hyperthermia) [7-12], nerves (Charcot Marie Tooth (CMT) [13; 14], motor neurone disease (MND) [15-19] hereditary spastic paraplegia [20-24], spinal muscular atrophies [25-27]) and neuromuscular junctions (myasthenic syndromes). [28] Muscle weakness, wasting, fasciculation, cramps, numbness, respiratory and cranial nerve palsies are common features in several NMDs. These disorders are frequently inherited and extremely heterogeneous with more than 500 implicated genes. There are many subgroups but individually they are rare and often severe, affecting a wide age range from children to adults. [29-34]

The advance in gene discovery can be attributed to progress in technology development and affordability of next-generation sequencing (NGS); moving from PCR in the 1980s to exome (protein-coding sequences) and gene panels in 2009/10 to the emerging use of whole genome sequencing in 2015[35]. With size and coverage come challenges with bioinformatics and variant interpretation. Given the extreme heterogeneity of NMDs and the large number of disease genes, they are well suited for the use of NGS[29]. Here, I also discuss the challenges of NGS and how the introduction of NGS into research and diagnostic clinical practice has changed the management of NMD patients.

1.2 Research and Diagnostic Challenges of NGS

NMD genetic heterogeneity is present in the group as a whole, as well as sub-groups. For example, in autosomal recessive limb girdle muscular dystrophy there are over 20 genes implicated and few diagnostic clues to narrow the genetic causes down. The aim of genetic testing is to make a molecular diagnosis but genetic analysis alone is not sufficient to prove causality of a particular genetic variant, the skill of the referring clinician is essential in assessing variants against the clinical differential diagnosis. A diagnosis of an inherited NMD may seem clear in the context of a positive family history or when multiple siblings are affected but the mode of inheritance is often difficult to define in small families or where family history may be lacking, as in apparently 'sporadic' cases where the mode of inheritance can be any pattern or *de-novo* dominant. The lack of family history and indistinct NMD phenotypes makes filtering of NGS data difficult, often requiring other unaffected family members to help with genetic proof and the need for further functional laboratory work as proof. [36; 37]

In the pre-NGS era, we and others have published screening and guidelines for genetic testing of NMD such as CMT [13] that were based on careful phenotyping and sequential Sanger sequencing of the most prevalent candidate genes. Although achieving a genetic diagnosis in over 60% of patients with CMT, Sanger sequencing of the ever-increasing number of individual genes is no longer cost or time effective. In CMT type 1, over 70% of patients have the 17p duplication and over 80% of patients receive a diagnosis using traditional methods but in CMT2, distal HMN and HSN, where 40–80% of disease-causing genes are unknown the diagnostic rates are only between 17 and 30%. [14] With the advancement in gene discovery and NGS,

we would expect these figures to be over 90% when all known and novel genetic causes could be analysed. [38]

1.3 Large NMD genes and different types of genetic defects

NMD genes are often large and contain many polymorphisms making the identification of disease-causing variants difficult. This is particularly true for myopathies where genes such as Duchene muscular dystrophy (DMD) (MIM#300377) spanning more than 2.3 Mb with 79 exons, the titin gene (TTN) (MIM#188840) with 363 exons and an open reading frame spanning over 100 Kb and nebulin gene (NEB) (MIM #161650) with 183 exons may be defective. Different types of defects are detected in NMD genes. [29] For example, 60% of patients with Duchenne and Becker muscular dystrophies have deletions in DMD, 10% have duplications, and the rest have point mutations or small insertions–deletions. Prior to NGS, Sanger sequencing, multiple ligation probe analysis (MLPA) and cytogenetic genomic hybridization (CGH) arrays would often be used to target mutation or deletion hot spots and therefore using unbiased NGS we will likely increase the frequency of detection of defects in these large genes. Although genome sequencing is capable of identifying deletions and duplications, the frequency of these defects in Duchenne and Becker muscular dystrophies may mean that MLPA/CGH array is still an easier and more cost-effective approach.

1.4 Next Generation sequencing

There are a number of challenges in moving from simple Sanger sequencing to the NGS analysis of multiple and eventually all the genes in the genome, aimed at

revealing the vast majority of genetic variants. The advent of NGS has in the last decade led to an era of inexpensive, high-throughput DNA sequencing of large numbers of genes in a single reaction, thus facilitating the discovery of novel disease genes and variants. [12; 16; 39-41]

Gene panels use next-generation sequencing to target genes and regions thought to be relevant to various diseases. Both custom-made panels that can include several hundred genes of interest and off-the-shelf panels of over 6000 genes are available. A custom-made panel with around 160 genes is used by the Diagnostics lab at the Institute of Neurology, and the same panel can be reduced to 40 genes when investigating neuropathies as well as when detecting mitochondrial deletions. Off-the-shelf panels such as Agilent focused exome and the Illumina Trusight panels offer larger-scale screening of genes that are thought important such as all those on *the Online Mendelian Inheritance in Man (OMIM)* database. They rely on solution hybridization, in which pools of biotinylated oligonucleotide probes, each targeting a segment of the desired capture region, are used to pull down particular pieces of the genome. Gene panels are a quick relatively low-cost screening method to look for a few gene variants associated with specific NMDs.

NGS technologies have also advanced the genetic diagnosis of mitochondrial disorders in sequencing the mitochondrial genome or nuclear genes. [42] A two-step strategy based on a targeted custom-made mitochondrial gene panel with 132 genes, and WES, was used by Legati *et al.* in a cohort of 125 early onset mitochondrial cases.[43] However, an uncontroversial genetic diagnosis was only

possible in 19 patients by targeted gene panel sequencing and in 6 out of 10 by WES. The rest of the cases although initially classified as mitochondria-disease patients, might be carriers of non-mitochondrial gene mutations.[40]

Targeted NGS can be beneficial for diagnosing patients with early presentation of NMDs. Chae *et al.* used targeted NGS to screen 43 patients presented with early onset neuromuscular disorders of an unknown genetic origin. They were tested by panel for 579 nuclear genes associated with myopathy.[29] Sanger sequencing has been the basic tool used for detecting mutations in myopathies but this is a laborious and expensive technique to be used for large genes. Compared to exome sequencing which is often costlier, has a higher false positive rate and turnaround time, and can be more difficult to interpret, targeted NGS is an attractive solution. Chae *et al.* were able to diagnose 21 of the 43 patients with a definite genetic cause for myopathy. [29] In a similar way, Evila *et al.* used a custom-made panel, MyoCap, with 180 myopathy related genes, and succeeded in diagnosing 21 out of 61 patients, with 9 of them having potential disease-causing mutations in TTN. [9]

Whole exome sequencing (WES) is an efficient technology that can increase the diagnostic yield when searching for alleles causing rare Mendelian disorders. [44] [45] Exome analyses the protein's encoding region where an estimated 85% of disease-causing mutations are believed to occur. Exomes are enriched by targeted hybridization and then each enriched genomic library is sequenced on a platform such as the Illumina Novaseq. This procedure covers around 96% of the targeted, mappable bases comprising the exomes of affected individuals. Filtering

using databases such as dbSNP and 1000 genomes removes common variants, aiding in identifying the causal gene (Figures 2 and 3). It has been an invaluable tool in gene discovery, with around 100 genes linked with neuromuscular disorders discovered in this way [46] such as the identification of ADCK3 mutations using exome sequencing.[47] However, less-reachable regions of the genome with low sequence complexity restrict the ability to design useful WES capture baits, and differences in the hybridization efficiency of WES capture probes can result in off target capture effects as well as regions of the genome such as in exon 1, or GC-rich regions, with little or no coverage. [39; 45; 48-50]

Whole genome sequencing (WGS) allows examination of single-nucleotide variants (SNVs), indels and duplications of more than 20-30bp, structural genomic variants (SVs) and copy number variants (CNVs) in both coding and non-coding regions of the genome. Protein-coding genes represent only 2% of the genome and most disease-associated variants are mostly found in these genes counting for at least 80% of Mendelian diseases. [51; 52] Therefore, WGS has more reliable sequence coverage with more uniformity and without the dependency to PCR. With the start of the 100,000 Project by the NHS in 2012, UCL has become a Genomics Medical Centre where patients with rare, life threatening and debilitating diseases such as many NMDs will be sequenced. Genomics England Clinical Interpretation Partnership (GeCIP) has also been established and aims to enhance clinical interpretation and validation of whole genome sequencing data delivered by the 100,000 genomes project. It is likely to reveal many new variants and genes and derive new scientific and clinical findings. With the rapid drop in sequencing cost and the ability of WGS to rapidly produce large volumes of data,

it is becoming a powerful tool for genomic research. However, WGS has still not been introduced in daily routine diagnostics, as it still remains more expensive and is much more computationally demanding in terms of bioinformatic data processing, capacity, often limiting its application in research frameworks. WGS data analysis is also extremely difficult to interpret given the large number of non-coding variants (98% of the human genome) since predicting their downstream biological consequence is very challenging.

Being more cost-effective and generating consistently higher diagnostic yield, WES still remains a better option in identifying nucleotide variants in genes involved in monogenic diseases and is more widely used in research as well as diagnostic situations for diagnosing genetically heterogeneous disorders. However, WES still has a few setbacks which include difficulties picking up deletions and insertions (especially those >20-30 base pairs) and difficulties in the detection of repeat expansions because of the short-read libraries (~100bp).

NGS is based on the parallel sequencing of DNA molecules spatially separated in a flow cell. This allows processing of millions of sequence reads in a time and cost effective manner. Different NGS chemistries and technologies are currently available, differing in specific technical details with respect to cost, accuracy, read length and multiplexing capability.

1.5 NGS limitations and improvements for clinical use

As NGS is employed both in research and diagnostic laboratories, some issues need to be addressed. One challenge of NGS is the detection of different types of

mutations, especially repeat expansions or structural variations. Some of the most common NMDs are due to repeat expansions; these mutations may be missed is larger than the read length of 150bp, leading to false negative results. Increasing sequence coverage, longer reads, improving bioinformatics algorithms, software and novel sequencing technologies may solve this problem in order to propose an exclusion diagnosis. [53] Alternatively, a combination of methods can be used. Detection of structural variations will benefit from whole genome sequencing. [54-57] Another obstacle is the incomplete coverage of commercial exome capture libraries. Although new versions of capture kits are released consistently, none capture all coding parts of all genes. This issue can be addressed by targeted sequencing of regions of interest and improving the capturing process.

Another drawback is the high error rates (false positives) in NGS compared to Sanger sequencing. Artificial mutations can be produced during template amplification or sequencing, leading to false positive results. Thus, Sanger sequencing of interesting variants detected by NGS is essential in validation but increases cost and turn-around time. This issue can be addressed by improving capture and sequencing approaches to increase variant coverage, achieving more reliable data. In addition, better data filtering protocols can reduce the pool of false positives. [53]

NGS generates a high volume of data which becomes problematic for analysis and storage. Cloud computing can be a solution for reducing the cost of expensive computing infrastructure but needs to be secure. Importantly, decreasing costs in NGS has outpaced the increase in calculation power and

storage capacity of computers. As sequencing becomes cheaper than data storage of the corresponding sequence output, re-sequencing of a patient DNA might be more cost-effective than saving original data. [58]

Large numbers of variants are detected in post-NGS filtering thus making it difficult to distinguish between individual, rare and non-pathogenic variations without clinical significance versus disease-causing mutations. Defining the complete list of polymorphisms in different populations will require sequencing of world population groups. A recent study hypothesised that 27 % of published mutations appear to be sequencing errors, common polymorphisms, or have a lack of evidence of pathogenicity. [55] This issue can be addressed by further analysis and validation such as *in-silico* prediction, detection of additional mutations in the same gene in unrelated individuals with the same disorder, absence in a control population, co-segregation in affected families and functional studies. Multiple criteria should be combined in order to prove the pathogenicity of the variation such as those provided by the Association for Clinical Genomic Science (Supplementary Figure 5). In addition, incidental findings can be problematic, especially for unreported diseases. Guidelines should be established and implemented by national committees, as in the case of the Genomics England sequencing project in order to determine whether the analysis and reports should be selective or complete. [36]

Identification of disease-causing mutations is vital for patients and families to provide an accurate and definitive molecular diagnosis. The knowledge of the mutation and mutated gene improves disease management, and allows for

inclusion into therapeutic trials. Genetic counselling is then possible, as carrier status determination and prenatal diagnosis can decrease the risk of recurrence [59]. In addition, finding the disease-causing mutation permits potential phenotype–genotype correlations and greater understanding of the pathophysiological mechanisms, disease pathways and development of potential therapies. [60; 61]

There are limitations and challenges in using NGS but this technology is significantly developed for research and routine diagnostic practice. As the cost of NGS declines and the technology improves leading to obtaining faster and more accurate data, it is strongly anticipated that this recent technology will complement many clinical and pathological investigations to enable diagnosis as well as the identification of disease modifying and pharmacogenomics factors.

KEY POINTS

- Neuromuscular disorders are clinically heterogeneous. There is a broad spectrum of disease ranging from myopathy and CMT through to the more aggressive neuromuscular impairment in MND.
- Next-generation sequencing technologies have helped greatly in diagnosing these disorders and underpinning the genetic pathways.
- Technology has evolved to create large disease gene panels through to whole exome sequencing, and final transition to whole genome sequencing.
- The understanding of genetic causes and disease pathways in the remainder of neuromuscular disorders will be key to identifying treatments.
- The formation of collaborative groups such as the neurology and neuromuscular GeCIP will greatly advance the interpretation and functional understanding of whole genome sequencing.

Thesis aims and outline

Genes and the way genes interact can play a pivotal role in the pathogenesis of rare and complex neurological disorders. Genetic changes can cause, predispose, protect, predict the course, direct the route of treatment and act as biomarkers. The scientific knowledge of neurogeneticists is exponentially increasing and understanding the pathophysiology of different inherited disorders has become key for providing medical care.

The overall aim of this thesis is to investigate the genetic architecture of rare complex neurodevelopmental disorders that fall under the umbrella of synaptopathies and are caused by synaptic gene dysfunction; both through identifying new genetic causes, as well as investigating new methods for validating genetic variants. This is focused on the following aspects:

1. To use whole-exome sequencing (WES) to identify the disease-causing gene or genes for a group of patients collected as part of the SYNAPS project. This will provide a genetic diagnosis for many of the patients, and also contribute to expanding the clinical and genetic spectrum of these conditions. This will focus mainly on the attempt to uncover the genetic basis of:

- an undiagnosed family with a central and peripheral demyelinating neuropathy, and thus establish the benefit of WES as a method for genetically diagnosing complex neuropathy patients.

- undiagnosed families with complex neurodevelopmental phenotypes, mainly with epilepsy and neuropathy and aim to replicate the findings in a bigger cohort.

2 To investigate the frequency of identifying neurodevelopmental gene mutations in a large synaptopathies case series. This will ascertain the degree of phenotypic and genetic overlap amongst these diseases, and help inform future genetic diagnosis protocols in these patients. Additionally, it may improve the limited understanding of the aetiology of these diseases.

I have provided a statement of contribution for each chapter and I have stated when work has been performed in collaboration with other colleagues or researchers.

Chapter 2. Materials & Methods

Mendelian disorders are classified as those with either one (heterozygous) or two (compound heterozygous or homozygous) mutations in a single genetic locus[62]. Current NGS platforms generate billions of bases of accurate nucleotide sequence data in short reads using reversible sequencing chemistries, thus expanding our ability to identify single nucleotide variations underlying Mendelian disorders[63]. NGS-based technique is the main approach in my project with the primary aim of identifying mutations in novel disease genes and elucidating molecular pathways involved in synaptic transmission disorders. These synaptic transmission disorders also known as synaptopathies include Mendelian diseases with paroxysmal neurological features or neurodevelopmental disorders characterised by the variable combination of intellectual disability, abnormal motor and language development features.

Synaptopathies are a group of diseases of the brain, spinal cord or peripheral nervous system relating to the dysfunction of synapses. Synapses are important components of neurons which allow information to travel co-ordinately throughout the nervous system to adjust behaviour to environmental stimuli and to control body functions, memories, and emotions. Optimal synaptic communication is vital for brain physiology, and any slight perturbations of synapse function can lead to brain disorders.

In recent years, research has demonstrated the relevance of synapse dysfunction as a major determinant of many neurological diseases. For example, synaptopathies caused by ion channel mutations are known as synaptic channelopathies and

episodic ataxia is an example. Abnormal synapse structure, function or physiology may contribute to neurodevelopmental disorders (autism, startle disease, and epilepsy) as well as neurodegenerative disorders (Alzheimer and Parkinson disease) (Figure 1). Some of these disorders will be studied further in this thesis.

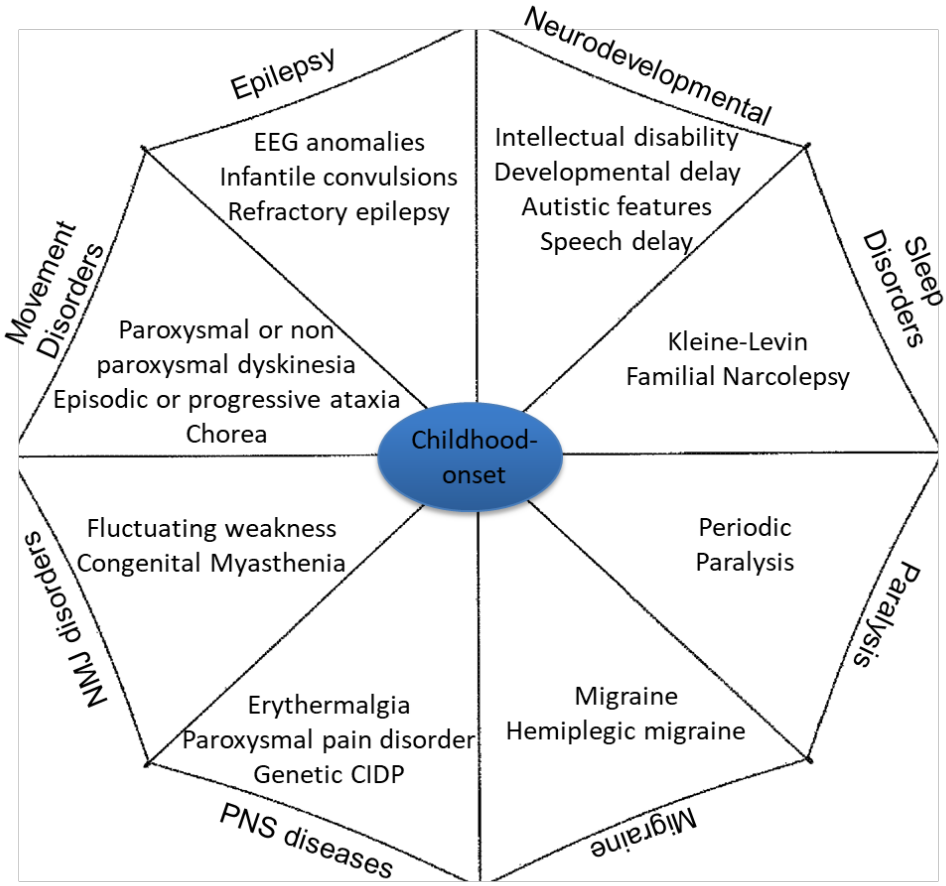


Figure 1. Spectrum of Synaptopathies or synaptic transmission disorders. A single umbrella made up of quite diverse pathologies; epileptic disorders, sleep disorders, migraine, peripheral and central neurodevelopmental diseases as well as neuromuscular junction disorders and paralysis.

2.1 The Synaptopathies and Paroxysmal Syndromes (SYNaPS) Study Group

As part of the Synaptopathies project grant awarded by the Wellcome Trust, patients diagnosed with paroxysmal neurological disorders and referred to the National Hospital for Neurology (in the period 2000-2015), who previously consented to donate their blood for research to the Neurogenetics Unit-Biobank at the UCL Institute of Neurology, were included. We further recruited patients and families by establishing the international research network "Synaptopathies and Paroxysmal Syndromes" (SYNaPS) Study Group (http://neurogenetics.co.uk/synaptopathies-synaps) to involve other clinicians and researchers worldwide. The main disease area of interest was episodic neurological disorders and synaptopathies associated with neurodevelopmental disorders.

2.1.1 Patient Recruitment

Between 2016 and 2018, around 4,000 individuals with different early onset episodic neurological and neurodevelopmental disorders were collected by the SYNaPS Study group, which now includes more than 50 clinical collaborators from more than 30 worldwide countries (Table 1 illustrates collaborators that have contributed with more cases). Recruitment of patients from genetically isolated regions such as Sicily, Cyprus and the Canary islands, proved more interesting as these population cohorts are enriched for founder mutations such as the *PRUNE*, *NKX6-2* and *PDE2A* variants identified in the past recent years by the Houlden Lab and published collaboratively[64-66]. Similarly, population cohorts in the Greater Middle East and South Asia (in countries such as Egypt, Jordan, Pakistan and Bangladesh) are enriched for recessive genetic traits, due to consanguinity.

Recruitment included patients and families presenting with neurodevelopmental disorders (such as delay in reaching motor milestones and speech, abnormal cognition and behaviour including psychiatric disturbances, autistic features and history of developmental regression) often associated with epilepsy and movement disorders. Clinical information such as NCS, EEG, MRI and other relevant investigations were collected to facilitate the analysis of possible disease-causing variants. All patients and families involved in the study have consented to be included in the research (ethics number UCL 07/Q0512/26), and institutions involved in the SYNaPS Study Group approved the research project collaboration. Blood (or occasionally saliva) samples were collected from participating patients and families and genomic DNA was extracted.

Table 1. SYNaPS Study Group collaborators worldwide.

Prof. Yamna Kriouile Prof. Mohamed El Khorassani	University of Rabat, Morocco University of Rabat, Morocco
Dr. Blagovesta Karashova Prof. Daniela Avdjieva	Medical University of Sofia, Bulgaria Medical University of Sofia, Bulgaria
Prof. Lionel Van Maldergem	University Hospital Liege, Liege, Belgium
Prof. Wolfgang Nachbauer	Medical University Innsbruck, Innsbruck, Austria
Dr. Larissa Arning	Ruhr-University Bochum, Bochum, Germany
Prof. Dagmar Timmann	University of Essen, Essen, Germany
Dr. Belen Pérez-Dueñas	University Hospital Sant Joan de Deu, Barcelona, Spain
Dr. Gabriella Di Rosa	University of Messina, Messina, Italy
Prof. Jatinder S. Goraya	Dayanand Medical College & Hospital, Ludhiana, India
Prof. Tipu Sultan	Children's Hospital of Lahore, Lahore, Pakistan
Dr. Salman Kirmani Prof. Shahnaz Ibrahim	Aga Khan University Hospital, Karachi, Pakistan Aga Khan University Hospital, Karachi, Pakistan
Prof. Jun Mine	Shimane University School of Medicine, Izumo, Japan

Prof. Hamed Sherifa	Assiut University Hospital, Assiut, Egypt
Prof. Selina Banu	University of Dhaka, Dhaka, Bangladesh
Dr. Issam Al-Khawaja	Albashir University Hospital, Amman, Jordan
Dr. Fuad Al-Mutairi	King Saud University, Riyadh, Saudi Arabia
Prof. Pierangelo Veggiotti	University of Milan, Milan, Italy
Prof. Yves A. Dauvilliers	University Hospital Montpellier, Montpellier, France
Prof. Michel D. Ferrari	Leiden University Medical Center, Leiden, Netherlands
Prof. Pasquale Striano	Institute “Giannina Gaslini”, Genova, Italy
Dr. Carmela Scuderi Dr. Eugenia Borgione	IRCCS OASI Maria SS. Troina, Troina, Italy IRCCS OASI Maria SS. Troina, Troina, Italy
Dr. Benigno Monteagudo Sanchez	University Hospital Arquitecto Marcide, Coruna, Spain
Prof. Alfons Macaya	University Hospital of Barcelona, Spain
Dr. Erin Torti	GenedX, Gaithersburg, Maryland, USA
Dr. Gali Heimer	University Hospital of Tel Aviv, Tel Aviv, Israel
Prof. Savvas Papacostas	Cyprus Institute of Neurology and Genetics, Cyprus

2.1.2 DNA extraction from blood samples

Genomic DNA (gDNA) extraction from blood samples was performed either in the Neurogenetics Laboratory (by the clinical diagnostic service team of the NHNN) or at LGC labs in Germany. This was performed using 5-10 ml of fresh or frozen whole blood samples using Flexigene kit (Qiagen, Germany) according to manufacturer’s instructions. The concentration and quality of DNA was measured using a NanoDrop ND-1000 spectrophotometer following the manufacturer’s instructions (NanoDrop Technologies, USA). Concentration was assessed at 260 nm. Purity was estimated by the 260/280 and 260/230 absorbance ratios, and the spectra of the ratios between 1.8-2.0 and 1.8-2.2 respectively were considered as DNA samples of the appropriate quality. The concentration of DNA was adjusted to appropriate values,

according to the technique to be used afterwards, by diluting the samples with autoclaved distilled H2O (dH2O) or TE buffer.

For the purposes of NGS library preparation, the measurement of DNA concentration required a higher precision than the one provided by the nanodrop. The Qubit® 2.0 Fluorometer from Thermo Fisher Scientific detects the amount of fluorescent dye bound to double stranded DNA, therefore providing a direct measurement of double stranded DNA quantity. The reagents used by the Qubit do not bind to degraded DNA or other molecules such as proteins, and are more accurate than the NanoDrop. DNA concentration of samples for NGS was measured with the Qubit and diluted with dH2O until reaching the required concentration for the library preparation used in each case. This usually required a few rounds of measurement/dilution.

For WES, DNA was diluted to 20 ng/µl or 50 ng/µl with autoclaved distilled H2O (dH2O) or TE buffer.

2.1.3 General methods of molecular genetics and next-generation sequencing analysis

Whole exome sequencing (WES) and Sanger segregation analysis were mainly used in this project to identify monogenic causes of neurological disorders and synaptopathies. For the generation of the data used in my project, WES was performed using a 100-bp or 150-bp paired-end reads protocol analysed on an Illumina HiSeq4000, Hiseq X or Novaseq platform. Sequence reads are aligned to the human reference genome (UCSC hg19). The samples were prepared according

to Agilent SureSelect Target Enrichment system version 6 (Figure 2) to capture genomic regions of interest (coding regions), enriching them out of an NGS genomic fragment library.

Library preparation is the preparation of genomic DNA that is to be sequenced. Each sample is sheared on a Covaris sonicator to generate fragmented double stranded DNA. Then the samples are transferred to a midi plate and incubated with AMPure XP beads on a magnetic stand. Repair ends are added to convert the overhangs resulting from sonication into blunt ends using ERP3 (End Repair Mix). The 3' to 5' exonuclease activity of this mix removes the 3' overhangs and the 5' to 3' polymerase activity fills in the 5' overhangs. Following end repair, the library size is selected using AMPure XP beads. A single 'A' nucleotide is added to the 3' ends of the blunt fragments to prevent them from ligating to each other during the adapter ligation reaction. A corresponding single 'T' nucleotide on the 3' end of the adapter provides a complementary overhang for ligating the adapter to the fragment. This strategy ensures a low rate of chimera formation. For this purpose, each sample is combined with A-tailing mix and a buffer. Ligate adapters ligate multiple indexing adapters to the ends of the DNA fragments, which prepares them for hybridization onto a flow cell. Each sample is then mixed with unique adapter indexes, ligation mix, and buffer so that samples can be then pooled together. After incubation with the adapters, a stop ligation buffer stops the reaction. The number of samples that are pooled together before sequencing will be dependent on the desired coverage. The more samples the lower coverage.

Enrichment uses PCR to amplify DNA fragments that have been ligated with adapters on each end of each molecule. PCR is performed with PPC (PCR Primer Cocktail) that anneals to the ends of the adapters. The amplified libraries are purified once again with AMPure XP beads. Library quantification can be performed with the Qubit, while library quality control is then performed on a Bioanalyzer (Agilent technologies).

Pooling and hybridization combine DNA libraries containing unique indexes into a single pool, and then binds targeted regions of the DNA with capture probes (exons and flanking regions in WES). Pooling the libraries requires specific amounts of each library according to the number of samples to make a single pool of 40 uL of 12 libraries. Then the library pool is combined with the CEX (Coding Exome Oligos) and a capture buffer, and loaded onto a thermal cycler. Streptavidin magnetic beads are used to capture probes hybridized to the targeted regions of interest. Two heated washes remove nonspecific binding from the beads. The exons enriched library is then eluted from the beads and prepared for a second round of hybridization.

The second hybridization binds targeted regions of the enriched DNA with capture probes a second time and ensures high specificity of the captured regions. The second round aims to capture probes hybridized to the targeted regions of interest. Two heated washes remove nonspecific binding from the beads. The enriched library is then eluted from the beads and prepared for sequencing.

Sequencing performed on an Illumina platform is a process of massive parallel sequencing by synthesis where all four nucleotides are added simultaneously to the

flow-cell channels, along with DNA polymerase, for incorporation into the oligo-primed cluster fragments. The DNA is linearized by cleaving one adapter and denatured to obtain single strands. Sequencing primers and four reversible terminators are added. The nucleotides carry a fluorescent label, and the 3'-OH group is chemically blocked so that each addition is a unique event. An imaging step follows each base incorporation step, during which each flow cell lane is imaged in three 100-tile segments by the instrument optics at a cluster density per tile of 30,000.

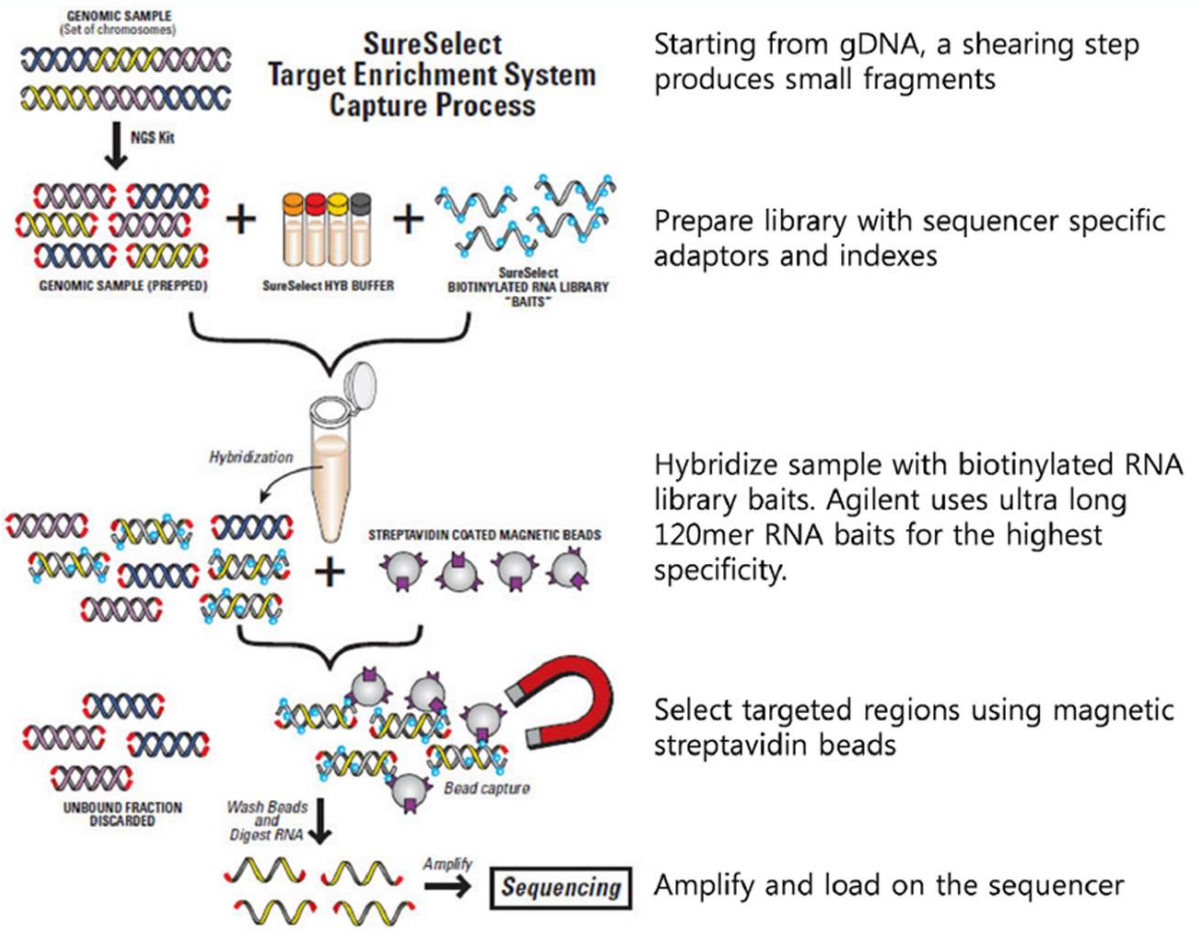


Figure 2. DNA library preparation and target enrichment. (Figure from Macrogen Inc.; https://dna.macrogen.com)

After each imaging step, the 3' blocking group is chemically removed to prepare each strand of DNA for the next incorporation of a nucleotide by the DNA polymerase. This series of steps continues for a specific number of cycles, as determined by user-defined instrument settings, which permits discrete read lengths of 25-100 bases (for example, 75 cycles will create seventy-five base pair sequencing reads. Most exons will be covered by 75-100 read lengths). The same sequencing process is performed from both DNA strand ends to create paired-end reads. This is particularly useful for accurate mapping and detection of structural variation. Data de-multiplexing is the next step to obtain files for alignment.

Target regions of the genome that present significant linkage in families can be a good strategy for variant investigation and filtering through WES. Linkage data analysed according to the pattern of inheritance can provide information on shared regions in affected family members compared with the unaffected relatives. In a similar way, homozygosity mapping is also useful in defining target regions in recessive families, where once the homozygous regions are delimited, the exome data is filtered to find the variants in those regions.

2.1.4 Bioinformatic processing of NGS data

The raw data was then processed by the Bioinformatics team at UCL (Dr. Alan Pittman, Hallgeir Jonvik and David Murphy) using a standardised bioinformatic pipeline with softwares such as bwa-0.7.12, GATKv3.4.0 or SnpEff, which allow the generation of all sequence reads with quality score assessment values for each read, alignment to the reference genome as well as annotated variants.

Paired-end sequencing reads in the form of FASTQ files are then aligned to the human genome reference (hg19) using Novoalign (Novocraft technologies). A BAM file is generated using SAMtools (http://samtools.sourceforge.net/). The removal of duplicate reads and the generation of statistics are performed with picard tools (http://picard.sourceforge.net). Local realignment of indels and variant calling was performed with the Genome Analysis Toolkit (GATK) https://software.broadinstitute.org/gatk/. Variant annotation was performed with ANNOVAR (http://www.openbioinformatics.org/annovar/). ANNOVAR is fully customizable and allows for annotation of the location of each variant (exonic, splice site, intronic, etc.), determination of its functional effect (non-synonymous change, stopgain, stoploss, frameshift, etc), adding population frequency from different databases (1000 genomes, Exac, EVS, dbSNP, etc), in silico predictions (Polyphen, mutation taster, CADD scores, etc), adding OMIM numbers, among other possibilities.

The types of files we generally obtain from NGS data are as follows:

1. Fastq: Human readable sequences with associated Phred scores. This is the native sequence or “raw data”. Fastq are usually paired-end. This means we get 2 files per sample (a forward and a reverse). The Phred score gives an estimate of the correctness of the corresponding base call.
2. SAM: Human readable mapped sequences, phred scores and coordinates to the reference sequence.
3. BAM: Binary version of SAM (compressed version of SAM). This file can be visualized with the Integrative Genomics Viewer

(http://software.broadinstitute.org/software/igv/) or the genome browser (http://goldenhelix.com/products/GenomeBrowse/).

4. VCF: Variant calling format. The VCF format is a tab delimited format for storing variant calls and individual genotypes. It can store both SNPs and indels.

5. Text files: After annotation with ANNOVAR the output can be retrieved in an CSV file that can be filtered with a text editor or also loaded onto Excel.

2.1.6 Variant prioritization

The above-mentioned pipeline generates a final list of variants which are then further filtered according to the inheritance/family pedigree being studied, the patients' phenotype and the genes of interest. In families where a recessive inheritance is suspected we usually look for homozygous or compound heterozygous variants, whereas in dominant disorders we would look for heterozygous variants. In X-linked disorders, we would be looking for variants in the X chromosome in males with absence of male-to-male transmission in the pedigree (Supplementary Figures 1-4).

Data from WES of a patient usually generates between 20,000-25,000 variants and only one (in the case of heterozygous variants associated with dominant inheritance or *de-novo* occurrence) or two (in the case of compound heterozygous or homozygous variants associated with recessive inheritance) will cause disease. Therefore, excluding benign, common and neutral variants as an initial step is very critical in the identification of a novel pathogenic gene variant in a Mendelian disorder [67; 68] (Figure 3). The population frequency of variants is a good way of prioritising variants. This is done by comparing the frequency of individual exome data variants against a wide array of large, publicly available population databases,

such as dbSNP, cg69, Exac and 1000 Genome project (www.internationalgenome.org). The rare diseases with episodic neurological symptoms or neurodevelopmental features described in this thesis are caused by mutations expected to be absent from population databases, or present at a very low frequency.

Quality scores (Phred scores), segmental duplication scores (refers to homology greater or equal to 96% meaning that these regions are likely to contain many false positive calls), and depth (number of reads at a specific location) are all important information provided by the analysis pipeline. They are vital scores used in the understanding if a variant is true or is possibly a false positive. NGS data is usually more reliable for single nucleotide changes than for indels.

The exome aggregation consortium (ExAC; exac.broadinstitute.org) database contains data from more than 60,000 adult individual exomes, but the diseases discussed in this work have onset typically in infancy or childhood and in certain cases are characterised by severe neurodevelopmental impairment. This therefore means that identifying healthy adult individuals carrying these rare variants will be very unlikely in such databases.

One of the biggest challenges in NGS interpretation is the frequent presence of sequencing artefacts, which can potentially lead to a misinterpretation of genetic data. Artefacts are usually not specific to the individual's exome, but nevertheless they seem to be often present in a large proportion of exomes (>1%) generated by the sequencing platform. They are generated by abnormal chemistry reactions

affecting the coverage and amplification of regions of the genome. It is very important to exclude variants present in >1% of the internal exome dataset to reduce the possibility of both artefacts and low-frequency variants, as well as confirm these by Sanger sequencing. Queen Square Genomics (set up in 2018) is the internal and in-house collection of human exomes in the department, which we can access and use to filter variants against. If a variant appears novel or rare in online databases but is common in our data from other disorders or controls it is more likely to be an artefact of our pipeline rather than a disease-related change.

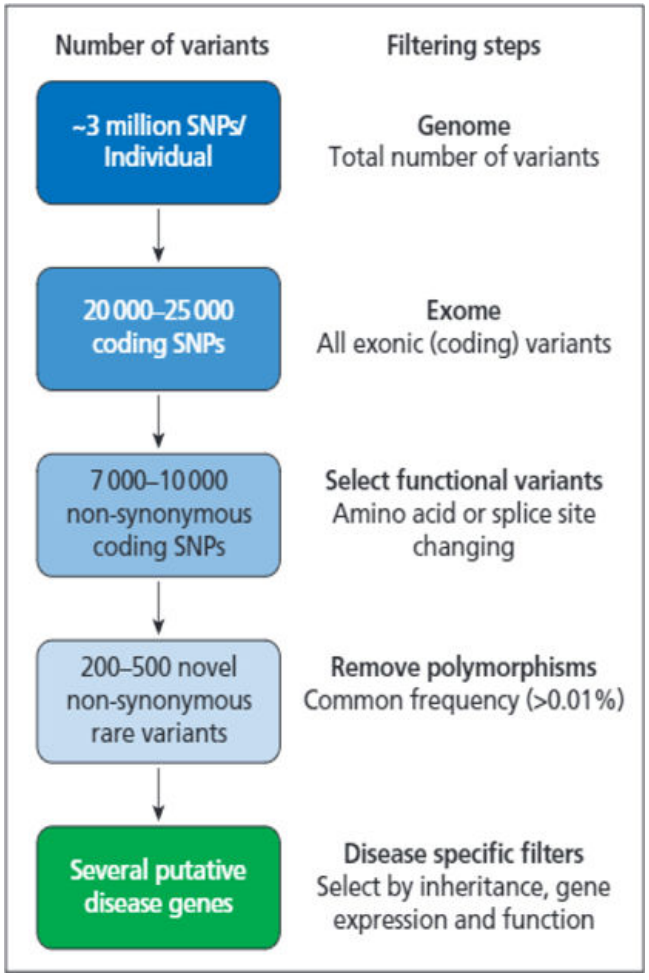


Figure 3. Filtering and functional analysis technique for whole exome sequencing. The human exome accounts for only around 1-2% of the human

genome. It is estimated that 85% of disease-causing mutations lie within the exome. In general, exome capture methods in NGS can give us output of data containing exons, splice-sites, UTRs, non-coding RNA sequences and intergenic regions. Data analysis of exomes involves the alignment of sequences, removing duplicates and annotating the data sequence and variants. Common polymorphisms and non-coding amino acid changes are removed to leave a list of unique or very rare (<0.01% of the population) variants that are either heterozygous or homozygous in investigating disease-causing mutations. While filtering for novel causes of disease, I focused on analysing exonic and splice-site variants. The number of variants usually obtained after this step is approximately 20,000.[69]

Where an autosomal dominant mode of inheritance is suspected, shared heterozygous variants among affected individuals are considered. *De-novo* dominant disorders can be investigated through sequencing of child-parents trios to identify *de-novo* heterozygous mutations; variants that are not present in both the mother and the father of the probands are considered *de-novo*. Recessive disease alleles can be present in normal population at a minor allele frequency (MAF) of < 0.05 or even <0.01 for rare diseases, even though these cut-offs are often subjective. In recessive disease filtering, homozygous, hemizygous or compound heterozygous variants are included. Missense, initiator codon, splice donor or acceptor, frameshift and stop lost/stop gained variants as well as in frame insertions or deletions are always included in the analysis.

In silico predictions are obtained with various prediction softwares which use different strategies to predict the pathogenicity of a specific variant. They use

physical and comparative properties, for instance if the variant is causing a truncating protein, if the amino acid change is to an amino acid with a different structure (such as proline) or hydrophilic/hydrophobic, if the change is located in a specific protein domain, if it changes the phosphorylation status of a protein, and/or the conservation among species. Evolutionary conservation is assessed with the Genetic evolutionary Rate Profiling ($GERP^{++}$) score (https://genome.ucsc.edu) and the possible impact of missense variants is evaluated according to their conservation across species [70; 71]. For a missense variant affecting a conserved residue, its $GERP^{++}$ must be >2 (arbitrary cut-off) to be associated with a Mendelian disease. The predicted deleteriousness of specific amino acid substitutions can also be assessed using *in silico* tools, such as CADD score (https://cadd.gs.washington.ed) to access coding and non-coding regions, SIFT (sift.bii.a-star.edu.sg) and Polyphen-2 (genetics.bwh.harvard.edu/pph2) which query non-synonymous changes and Mutation Taster (www.mutationtaster.org) which works for small indels.

To conclude, pathogenic variants are best narrowed down to those that: (i) segregate with the disease within a kindred (ii), change amino acid(s) or a splice site, (iii) are conserved across species, (iv) occur in genes that are expressed in CNS regions of interest according to their neurological phenotype (http://ukbec.wordpress.com/), (v) are predicted as damaging mutations based on *in silico* prediction tools (e.g. CADD/Polyphen), (vi) are present in brain channels or synaptic core genes or genes crucially involved in the vesicles release machinery, and (vii) have a frequency <1% in Exome Aggregation Consortium (ExAC) and an exome in-house database (containing >5,000 individuals). Variants present in >1% of our internal exome dataset at the UCL Institute of Neurology (containing ~ 10000

exomes from individuals affected with a range of neurological disorders as well as exomes from healthy control individuals from different populations and ethnic groups) are excluded. Based on values from the ExAC database we prioritise variants in genes with high probability of being intolerant to loss of function (i.e. ExAC pLI >0.8) and intolerant / constrained for missense variations (ExAC z-score >2), variants affecting domains important to genes function (according to the InterPro database; https://www.ebi.ac.uk/interpro), variants in genes predominantly expressed in the central nervous system (Genotype Tissue Expression Consortium database; https://gtexportal.org), and variants in genes implicated in brain development and function particularly in key synaptic transmission processes (from literature). Synonymous variants are usually excluded during analysis as they do not result in amino acid substitutions.

2.1.5 Homozygosity mapping

SNP arrays were run by UCL Genomics, Microarray and High Throughput Sequencing Facility at the Institute of Child Health, using the Illumina Infinium Global Screening Array version 2 which contains probes for over 200,000 markers distributed across the genome. And it is processed in multiples of 24 arrays at a time. 10 µl of DNA at a concentration of ~75 ng/µl were prepared and samples were processed, hybridized and scanned in accordance with the manufacturer's instructions. SNP raw data was converted to VCF file by the bioinformatics team (as described before) and cross matched to the reference sequence. Data were analysed in PLINK v.1.9 to create plink files which can be then used to carry out haplotype analysis or homozygosity mapping.

Homozygosity mapping was carried out either using WES or SNP-array data and the online-tool “homozygosity mapper” (http://www.homozygositymapper.org specifying affected and unaffected family members. Genes in all stretches of homozygosity > 1Mb were cross-queried with a list of candidate genes filtered from post-filtering WES data and a list of neurodevelopmental conditions resembling our patients.

2.1.6 Sanger sequencing

Segregation of candidate dominant or recessive variants is done by traditional Sanger sequencing in other affected and unaffected family members and this always helps to narrow down the list of candidate variants (post-WES filtering). Sanger sequencing was first developed in 1977 by Frederick Sanger and became the main method for sequencing DNA until the development of next generation sequencing only a few years ago. Sanger sequencing, based on the chain-terminator method, still remains the gold standard method for mutation confirmation. This method has been extensively used in this thesis, and is based on DNA polymerase randomly inhibited by a small amount of modified dideoxynucleotides (ddNTPs) which are mixed with normal deoxyribonucleotides (dNTPs). This produces newly synthesize DNA fragments of different lengths with different terminators, which enables one subsequently determine its sequence by capillary electrophoresis.

The region of interest was amplified by polymerase chain reaction (PCR) using flanking primers. PCR amplification was confirmed in most cases by agarose gel electrophoresis and then PCR products were purified using ExoFAST (Thermofisher Scientific). Sequencing was performed in both directions using Big Dye Terminator (v.3.1, Thermofisher Scientific), and run on an ABI 3730 DNA analyser (Applied Biosystems). Results were analysed using Sequencher software (Gene Codes) and

were compared to the reference sequence. However, this approach can be challenging as in dominant Mendelian diseases an age dependent incomplete penetrance of the disease may occur, increasing the risk of identifying a disease variant in an apparently unaffected individual.

2.2 General methods of functional analysis

2.2.1 Primer design

Primers were designed using Primer3 (http://primer3.ut.ee/) as follows. I first input the nucleotide sequence from each exon separately in the program, along with approximately 500bp flanking sequence from each of the upstream and downstream intron. The DNA sequence was extracted from the suitable genonic sequence downloaded from the Ensembl website (http://www.ensembl.org/index.html). A suitable transcript is usually the longest coding transcript with a consensus CDS (CCDS) identifier (ID). Then we selected the target region that should definitely be included in the polymerase chain reaction (PCR) product (which is the exon and approximately 100bp flanking intronic regions). The primers were generated under the default parameters in Primer3. This software allows researchers to design specific primers for a region under specific conditions but for 1 region at a time. Optimum primer size is usually around 20 BP, with optimal melting temperatures set for between 55°C and 65°C, the primer GC content is set at around 50 % (30-70), and low self-complementarity is preferable (to avoid a large amount of primer dimers). I then blasted them on https://genome.ucsc.edu/cgi-bin/hgPcr to ensure they were specific for the target region, and do not bind unspecific sites, and finally on online databases to ensure there are no common SNPs within the primer sequence. Primer pairs overlapping known SNPs were discarded. Where an exon

was too long (over 600bp), I either designed primers for two separate amplicons or designed primers for the single amplicon with additional mid-amplicon primers for use in sequencing reaction. When two neighbouring exons were separated by a short intron, I designed primers amplifying them as a single amplicon. Finally, I ordered the designed primers from Sigma already resuspended in autoclaved, double distilled MilliQ water. The stock solutions were diluted to a final working concentration of 5-10 pmol to be used in PCR.

Primer sequences used in this thesis can be found in the Appendix.

2.2.2 Polymerase Chain Reaction (PCR)

All coding exonic regions of the studied genes were amplified by PCR and sequenced by Sanger sequencing (dideoxyterminator nucleotide reaction). First, primers were optimised using control DNA to ensure efficiency of the reaction, which was confirmed through agarose gel electrophoresis.

Table 2. Example of PCR mix recipe.

Reagent	**Quantity (μl)**
Roche fast start PCR master (Roche, UK)	7.5
Forward primer (10pM)	1
Reverse primer (10pM)	1
gDNA (~50ng/μl)	1
Autoclaved ddH_2O	5

PCR reactions were then loaded onto an Eppendorf Mastercycler thermal cycler. Different cycling programs can be used, usually the first try was using a touchdown

temperature according to the optimal annealing temperature of the primers (PCR program 65td55). The PCR programs all contained a first cycle of denaturation of the double-stranded DNA, then annealing of the primers, and a final elongation step adding each deoxynucleotide (dGTP, dCTP, dATP and dTTP). An example is shown on Table 3.

Table 3. PCR 65 touchdown 55 cycling programme.

Step	Temperature	Time	Number of cycles
Denaturation	94°C	10 minutes	X1
Denaturation	94°C	30 seconds	X8
Annealing	65°C	30 seconds	
Elongation	72°C	45 seconds	
Denaturation	94°C	30 seconds	X16
Annealing	65°C (-0.7°C per cycle)	30 seconds	
Elongation	72°C	45 seconds	
Denaturation	94°C	30 seconds	X16
Annealing	55°C	30 seconds	
Elongation	72°C	45 seconds	
Elongation	72°C	5 minutes	X1
Hold	4 °C		

If the bands on the agarose gel electrophoresis were satisfactory but faint, we increased the number of cycles. If there are multiple bands or no bands at all, we repeated the PCR reaction adding different concentrations of dimethyl sulfoxide (DMSO), Q-solution and/or Betaine solution (Sigma, UK). If none of these combinations worked, we tried different cycler programs and/or designed new primers.

2.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to determine whether the PCR was successful by visualising PCR products on a 1% agarose gel. We prepared 10x Tris/Borate/ Ethylenediaminetetraacetic acid (EDTA) (TBE) buffer according to the recipe described in table 4.

Table 4. 10x Tris/Borate/EDTA (TBE) buffer recipe.

Reagent	**Quantity**
Trizma base	121.1 g
Boric acid	61.8 g
Ethylenediaminetetraacetic acid (EDTA)	7.4 g
Autoclaved ddH_2O	To a final volume of 1L

*EDTA, Ethylenediaminetetraacetic acid; ddH2O, double distilled water.

10xTBE buffer was diluted to 1x using double distilled water (ddH_2O). To prepare a 1% agarose gel, 100 ml of 1xTBE were mixed with 2 g of agarose and heated in the microwave oven for 2:30min. When the solution was transparent, we added 10 µl of GelRed dye and mixed by shaking until the solution was homogeneous, to produce a

porous gel in which DNA fragments can move. This solution was then poured in a gel casting mould after adding the combs and taking care not to create bubbles, and was left to set.

We prepared 6x Orange G dye according to the recipe described in Table 5.

Table 5. 6x Orange G recipe. *ddH2O, double distilled water.

Reagent	Quantity
ddH_2O	4 ml
Glycerol	1 ml
Orange G powder (Sigma)	1 g

We mixed 3 µl from each PCR product with 3 µl of Orange G dye and loaded this mix for each sample in one well of the gel. In each lane of the gel, we added 5 µl of DNA and started the electrophoresis at 120V for 30min. When an electrical current is applied across the gel, the negatively charged DNA will travel at a rate relative to its size. Contamination screening was also performed by visualization of the same reaction containing dH_2O instead of DNA. A 1 kb DNA ladder (Qiagen) was used to judge the size of the amplified fragments and ensuring they matched the target region. The DNA bands were visualised under an Ultraviolet (UV) transilluminator, and digital photographs were taken using the Syngene GeneGenius image acquisition system and GeneSnap software (Synoptics).

2.2.4 PCR purification

PCR purification was done using the Exo-Sap method, by preparing a stock solution including 50 µl Exonuclease- I (Thermofisher scientific), 200 µl alkaline phosphatase

(Fast-AP, Thermofisher scientific) and 750 µl ddH2O. We added 2ul of this solution to 5 µl of PCR product and placed the plate in the thermocycler for an on the thermal cycler for 30 minutes at 37°C followed by 15 minutes at 80°C.

2.2.5 BigDye sequencing reaction

For the sequencing reaction, the BigDye terminator v1.3 Cycle sequencing kit (Applied Biosystems, USA) was used with the below recipe and loaded into the PCR cycler with the following program run: 1 cycle of denaturation at 94°C for 1 minute, followed by 25 cycles of: denaturation a 94°C for 30 seconds, annealing at 50°C for 15 seconds, and elongation at 60°C for 4 minutes.

Table 6. Sequencing reaction recipe.

Reagent	**Volume (µl)**
ABI sequencing buffer	2
Primer (forward or reverse) (10pM)	2
Autoclaved ddH_2O	1.5
BigDye	0.5
purified PCR product	5

For amplicons that we had used Q solution or DMSO in the PCR, this was also added in the sequencing reaction. Both forward and reverse were sequenced for confirmation.

2.2.6 Sequencing reaction purification

Sequencing products were purified using the Sephadex plates and by preparing a solution of 40 ml distilled, autoclaved H_2O and 2.9gr Sephadex in a 50 ml tube. We allowed this solution to hydrate at room temperature for 30min prior to usage. When we did not plan to use the solution immediately, we stored it at +4^oC but warmed to room temperature prior to usage. After mixing well, we added 350 μl of Sephadex solution per well of a Corning® glass filter plate (Corning® FiltrexTM 0.66 mm glass fibre filter, Sigma) placed on an empty collection plate. We then prepared a balance of equal mass and centrifuged them at 750xg for 3min. The plate does not need to have a lid when centrifuging. We then diluted the sequencing reaction product (~10 μl) with ddH_2O to a final volume of 20 μl and pipetted it onto the Corning plate that is now placed over a non-skirted plate. When pipetting, we were careful not to touch the tips on the Sephadex solution in the Corning plate. The plate was then centrifuged at 910xg for 5min, producing 20 μl of purified sequencing reaction product in the non-skirted collection plate. We added ddH_2O to any empty wells of the non-skirted plate prior to placing in the sequencing machine. Each Corning plate can be used up to five times, after cleaning it as follows: We dispose of the Sephadex columns, rinse the plate 4 times with ddH_2O and centrifuge at 910xg for 3min to remove the residual water. Sequencing reaction products were read by the ABI 3730xl DNA analyser (Applied Biosystems, USA), which is a capillary electrophoresis system.

2.2.7 Variant segregation

The sequencing results were analysed using the Sequencher software version 4.1.4 by comparing each sample's sequencing against the reference of the gene taken

from the Ensembl or NCBI website. The sequences were checked for single nucleotide changes and insertions and deletions in coding and flanking regions of the target gene. Sequence variation was checked to see if it was previously reported and at which frequency utilizing online databases including Ensembl, the exome variant server (http://evs.gs.washington.edu/EVS), dbSNP (http://www.ncbi.nlm.nih.gov/SNP/), the 1000genomes project (http://browser.1000genomes.org/index.html), and the Exome Aggregation Consortium (http://exac.broadinstitute.org/).

An example of the output of Sanger sequencing is presented in Figure 4.

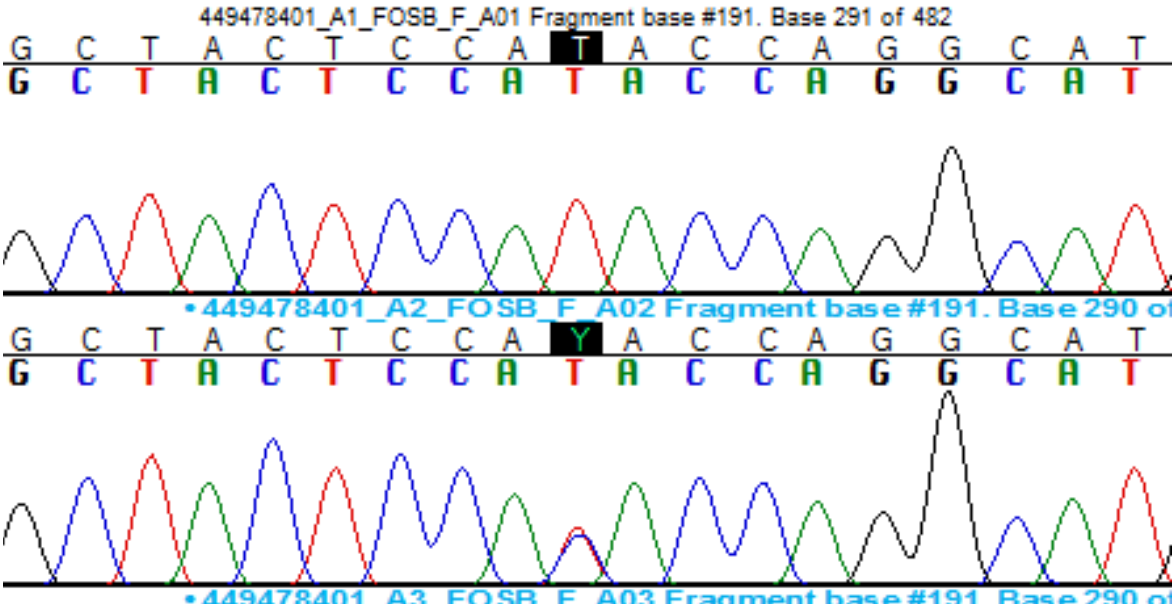


Figure 4. Sanger sequencing chromatogram result of a section of the gene *FOSB* exhibiting reference sequence in the top sample and a homozygous change in the bottom patient sample.

2.2.8 General cell culture procedures

Primary skin fibroblast cultures were obtained from healthy controls and patients consenting to a skin biopsy. Primary fibroblasts were cultured in Dulbecco's Modified

Eagle's Medium (DMEM, Life Technologies) supplemented with 10% fetal bovine serum (FBS, Hyclone), 1% penicillin and streptomycin (PS, Life Technologies) at 37°C in a 5% CO2 atmosphere. All cells were screened for mycoplasma using LookOut Mycoplasma PCR Detection Kit (Sigma) on a routine basis.

Primary lymphoblast cultures were also obtained from patients in ACD yellow-top tubes and were processed at ECACC (Salisbury, UK) where lymphoblastoid cell lines (LCLs) are routinely generated by Epstein - Barr virus (EBV) transformation of the B-lymphocyte component within the peripheral blood lymphocyte (PBL) population of a blood sample. Once a cell line has been successfully established, a bank of four ampoules is generated and stored in the ECACC cryostorage facility which our lab can request for cell culturing. Both fibroblasts and lymphoblasts can be used in Western blot protein expression or imaging experiments.

2.2.9 RNA Extraction

RNA from cultured fibroblast or lymphoblasts was extracted by using the Zymo Quick-DNA/RNA Miniprep Plus Kit (D7003) by following the below steps as outlined in the manufacturer's protocol:

- Pellet mammalian cells by centrifugation (≤500 x g for 1 minute), remove the supernatant and resuspend the cell pellet in DNA/RNA Lysis Buffer (300 µl buffer for ≤5x10^6 of cells)
- Transfer the sample into a Spin-Away™ Filter (yellow) in a Collection Tube and centrifuge.
- RNA is in the flow-through. Transfer the Spin-Away™ Filter (yellow) into a new Collection Tube and add 1 volume ethanol (95-100%) to the flow-through

and mix well. Then transfer the sample into a Zymo-Spin™ IIICG Column (green) in a Collection Tube and centrifuge. Discard the flow through.

- Add 400 µl DNA/RNA Prep Buffer to the column and centrifuge. Discard the flowthrough.
- Add 700 µl DNA/RNA Wash Buffer and centrifuge. Discard the flow-through.
- Add 400 µl DNA/RNA Wash Buffer and centrifuge the column for 2 minutes to ensure complete removal of the wash buffer. Carefully transfer the column into a clean microcentrifuge tube.
- Add 100 µl DNase/RNase-Free Water directly to the column matrix, let stand for 5 minutes, and then centrifuge to elute DNA and RNA from the respective column. The eluted DNA & RNA can be used immediately or stored at ≤-70ºC.

2.2.10 Reverse transcription (RT-PCR)

All samples were handled on ice at all times. RNA was quantified and quality checked using a Nanodrop and/or Agilent 2100 bioanalyzer and stored away at -80°C or processed immediately for cDNA generation. To generate complementary DNA, reverse transcription with 1 µg RNA-input using SuperScript® III First Strand Synthesis System from Invitrogen (18080051) was performed for all experiments mentioned in this thesis.

In each reaction, reagents from Table 7 were mixed and incubated for 5 mins at 70°C, then immediately put on ice for 5 mins, before components of Table 8 were added, the reactions spun down and cycling conditions detailed in Table 9 applied. From each batch of extractions, one separate reaction was added as a negative

control where everything was done as with all other samples apart from water was added instead of the enzyme Reverse Transcriptase.

Table 7. Reverse transcription using SuperScript III step 1.

Reagent	**Volume per reaction (µl)**
Random primers 3 µg/µl (Invitrogen 48190011)	1
dNTPs, 10 mM	1
1 µg RNA + ddH2O, RNAse free	11
Total	**13**

Table 8. Reverse transcription using SuperScript III step 2.

Reagent	**Volume per reaction (µl)**
Total from step 1	13
DTT (Invitrogen)	1
First Strand buffer 5x	4
RNase OUT (Invitrogen)	1
RT+: SuperScript III RT; or RT-: ddH2O, RNAse free	1
Total	**20**

Table 9. RT-PCR cycling conditions SuperScript III.

	RT-PCR SuperScript III	
Temperature (C°)	**Time (min)**	**Number of cycles**
25	05:00	1
42	10:00	1
50	30:00	1
70	15:00	1
4	Hold	-

Obtained 20 µl cDNA (concentration 50 ng/µl) was diluted 1/10 using 180 µl of RNAse free ddH2O to yield 200 µl cDNA (concentration 5 ng/µl) which was used as input for subsequent qPCR.

2.2.11 Semi quantitative PCR

For all quantitative PCR experiments described in this thesis, exon-exon spanning primers were designed, which are specific for the amplification of cDNA, as they can only bind to the exon sequences of the gene of interest and not to genomic DNA which would contain introns. 5 µg of cDNA was determined as input for qPCR using GoTaq® G2 Green Master Mix in the recipe described in Table 10 and run using the protocol in Table 11. For all experiments, one RT^- sample, and 1 sample where water is added instead of cDNA were run for each primer pair on each semi-qPCR run.

Table 10. Semi-qPCR master mix.

Reagent	Volume per reaction (µl)
GoTaq® G2 Green Master Mix	12.5
ddH2O, RNAse free	10
Primer for (c= 10µM)	1
Primer rev (c= 10µM)	1
Input cDNA (10 ng/µl)	0.5
Total	**25**

Table 11. Semi-qPCR cycling conditions.

	Semi-qPCR	
Temperature (C°)	**Time (min)**	**Number of cycles**
94	10:00	1x
94	00:30	32x
65	00:30	
72	00:45	
72	05:00	1x
4	Hold	

Methods of Chapter 3.2

Molecular Biology

NFASC missense constructs were cDNA synthesised by Genescript (USA) into pCMV-3Tag-4A by replacing the coding regions (CDs) for these inserts via using restriction enzymes BamHI and XhoI (NEB), giving rise to all NFASC plasmid variants (Figure 5).

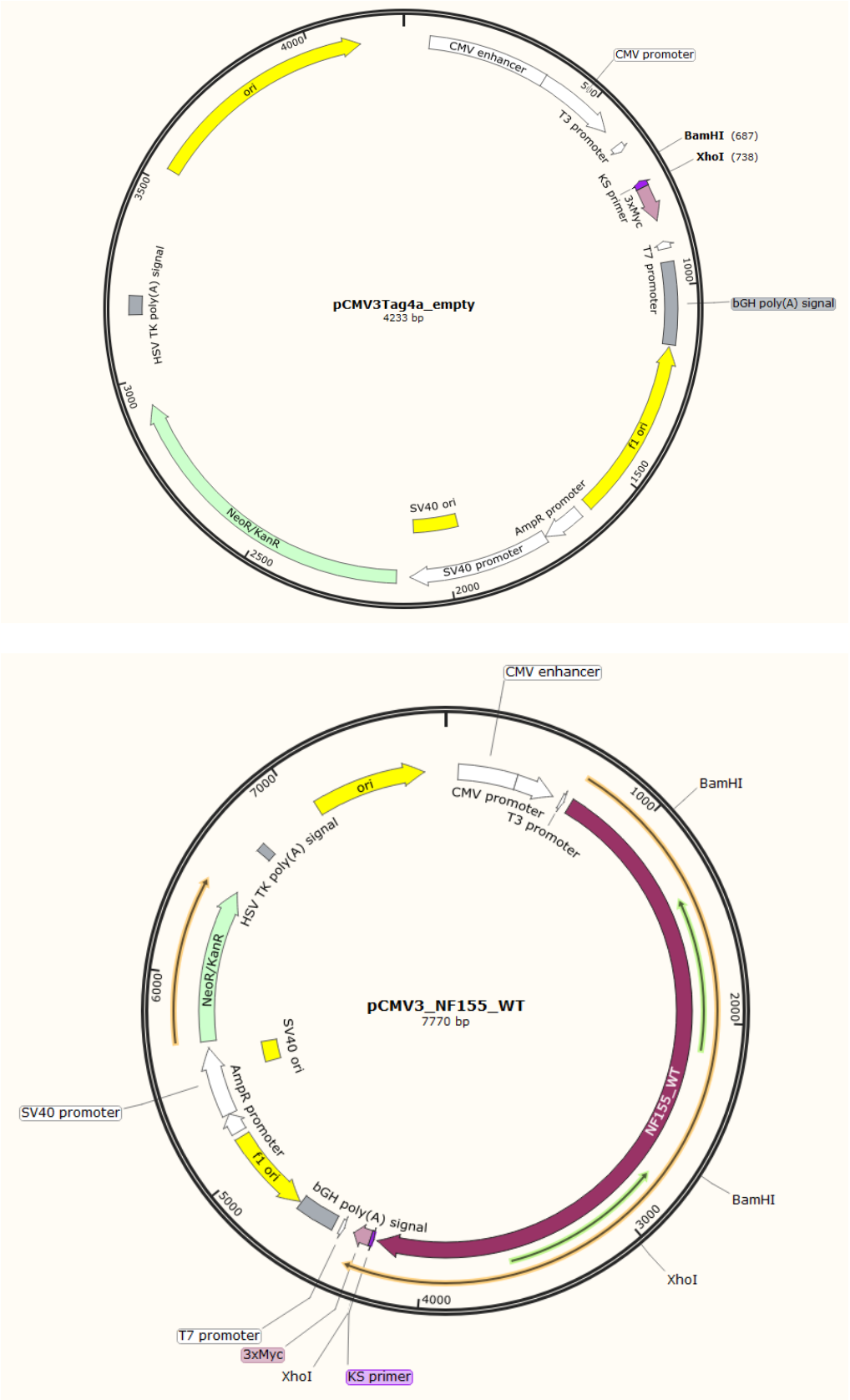


Figure 5. Plasmid maps of the NFASC construct using Snapgene software. An empty pCMV-3Tag-4A plasmid (top) was used as a template for all NFASC isotope

plasmids. An example plasmid with a NFASC155 wild-type (bottom) as generated by Genscript (USA) using restriction enzymes BamHI and XhoI (NEB) to replace the CDs for the NFASC155 WT insert. The same was generated for NFASC186 and NFASC3.

The variants generated were pCMV_Nfasc186[WT], pCMV_Nfasc186[N130D], pCMV_Nfasc186[R359P], pCMV_Nfasc186[P694T], pCMV_Nfasc186[S820P], pCMV_Nfasc155[WT], pCMV_Nfasc155[N124D], pCMV_Nfasc155[R370P], pCMV_Nfasc155[P705T], pCMV_Nfasc155[S831P] and pCMV_Nfasc155[P939Ter]. To generate pCMV_Nfasc155[P939Ter], the respective variant was introduced with the QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies) as outlined below in table 12. A Myc epitope was then inserted before the stop codon into pCMV_Nfasc155[P939Ter] using site-directed mutagenesis with primers: 5-GAACAAAAACTCATCTCAGAAGAGGATCTGTGATTGGATACACTCTCAAAT-3 and 5-CAGATCCTCTTCTGAGATGAGTTTTTGTTCTGATCCCATTTGGATGCTCAG-3.

Table 12. QuickChange Site-Directed reaction recipe.

Reagent	**Quantity (**µl)
10× reaction buffer	5
dsDNA template (5–50 ng)	4
Forward mutagenesis primer (125 ng)	1
Reverse mutagenesis primer (125 ng)	1
dNTP mix	1
PfuUltra HF DNA polymerase (2.5 U/µl)	1 (added last)
ddH2O	to a final volume of 50 µl

Each reaction was processed using the cycling parameters outlined in the following table.

Table 13. PCR conditions using the QuickChange Site-Directed mutagenesis kit.

Temperature (°C)	Time (min)	Cycles
95	00:30	1
95	00:30	12-18
55	01:00	
68	1minute/kb of plasmid	

1 µl of the Dpn I restriction enzyme (10 U/µl) was added at the end of the cycling program, mixed gently and immediately incubated at 37°C for 1 hour to digest the parental supercoiled dsDNA. 1 µl of the Dpn I-treated DNA from each control and sample reaction was transformed into separate 50-µl aliquots of XL1-Blue supercompetent cells using the below protocol:

1. Pre-chill two 14-ml BD Falcon polypropylene round-bottom tubes on ice. Preheat SOC medium to 42°C.
2. Thaw the supercompetent cells on ice. When thawed, gently mix and aliquot 100 µl of cells into each of the two pre-chilled tubes.
3. Add 1.7 µl of β-mercaptoethanol provided with this kit to each aliquot of cells.
4. Swirl the contents of the tubes gently. Incubate the cells on ice for 10 minutes, swirling gently every 2 minutes.
5. Add 20 ng of the experimental DNA to one aliquot of cells. Swirl the tubes gently.
6. Incubate the tubes on ice for 30 minutes.

7. Heat-pulse the tubes in a 42°C water bath for 45 seconds. The duration of the heat pulse is critical for maximum efficiency.

8. Incubate the tubes on ice for 2 minutes.

9. Add 0.9 ml of preheated SOC medium and incubate the tubes at 37°C for 1 hour with shaking at 225–250 rpm.

10. Plate ≤200 µl of the transformation mixture on LB agar plates containing kanamycin antibiotic.

11. Incubate the plates at 37°C overnight.

12. The number of colonies obtained will vary according to the size and form of the transforming DNA, with larger and non-supercoiled DNA producing fewer colonies.

The next day, colonies were picked and amplified using QIAprep® Spin Miniprep Kit according to the manufacturer's protocol which involves pelleting overnight bacterial cultures, and alkaline lysis of bacterial cells followed by adsorption of DNA onto silica in the presence of high salt. This resulted in isolation of up to 20 µg high-purity plasmid DNA which we could use further on in cell-based assays.

12 mm glass coverslips are flame sterilized and inserted into 24 well-plates (coverslips have previously been washed with 20% Nitric acid overnight, then water, and 95% ethanol). The wells are then are coated with 500 µl of poly-L-lysine (0.1 mg/mL) diluted in ddH2O for 2 hours at 37°C. Wells are then washed 3 times 5 minutes with ddH2O, and the wells are dried.

HEK 293 cells are grown in 75 cm^3 flasks. The medium is withdrawn and the cells are washed with 5 ml of PBS, then incubated with 2 ml of trypsin/EDTA 0.5 % for

5min at 37°C. The trypsination is stopped with 10 ml of culture medium (DMEM glutamax + 10% FBS + Non-essential amino acid + sodium pyruvate + Peni/Strepto), and the cells are suspended and centrifuged for 5 min at 500 g. The cell pellet is then re-suspended in culture medium and counted. Ideally, cells need to be 50-70% confluent on the day of transfection. The wells are seeded with HEK293 cells (around 40 000 cells/wells in 1 ml of media) and the day after, the cell medium is changed and the cells are transfected as follows.

Cell assays

Human embryonic kidney (HEK293) cells were transfected with Nfasc155 or Nfasc186 variants (1 µg of DNA in 150 mM NaCl to a final volume of 50 µl) using JetPEI Polyplus transfection agent (2 µl of jetPEITM in 150 mM NaCl to a final volume of 50 µl). After 48 hours, cells were incubated for 20 minutes at 37°C with rabbit antibodies against Nfasc186 (D6G60; cell signaling technology) or Nfasc155 (D7B6O; Cell Signaling Technology) diluted 1/50 in Opti-MEM or with 1 µg of gliomedin-Fc [72] diluted in Opti-MEM (ThermoFisher Scientific).

Cells were then washed in phosphate buffer saline (PBS) in 3 successive washes to remove any unbound antibodies, fixed with paraformaldehyde 2% in PBS and blocked with blocking solution (5% fish skin gelatin containing, 0.1% Triton X-100 in PBS). HEK293 cells were then incubated with a mouse anti-Myc (Roche; 1/200) for 1 hour, then incubated with Alexa conjugated secondary antibodies (1/500; Jackson ImmunoResearch) for 30 minutes. Coverslips were washed three times in PBS, stained with DAPI, and mounted with Mowiol plus 2% DABCO. In some experiments, HEK293 cells were co-transfected with GFP and Nfasc155 variants. In those

experiments, cells were fixed and revealed with a mouse anti-Myc. Immunostaining were examined using an ApoTome fluorescence microscope (ApoTome, AxioObserver and AxioCam MRm, Carl Zeiss MicroImaging GmbH). Digital images were manipulated into figures with CorelDraw and Corel Photo-Paint.

Cell aggregation assay

N2A cells were plated in 6-well plates at a density of 500,000 cells/well and were transiently transfected using JetPEI (Polyplus-transfection) with Nfasc155 and ptdTomato-N1 (Clontech), CNTN1/CASPR1 and peGFP-N1 (Clontech), or peGFP-N1 alone. Two mixtures were prepared and left at room temperature for 20 minutes, one of them being NaCl (100µl) with JetPEI (6 µl) or NaCl (100µl) with DNA (2.5 µg).

The arrangement of DNA plamids mixed and added to each well were as follows. They were left at 37°C for 24 hours.

	1	**2**	**3**	**4**	**5**	**6**
hNF155 (3.6 µg/µl)		0.56				
hNF155-R705P (1.5 µg/µl)				1.33		
hCNTN1 (1.1 µg/ml)					0.9	0.9
hCaspr1 (1.35 µg/ml)					0.74	0.74
GFP (0.6 µg/µl)	0.8				0.8	0.8
Tomato (1.9 µg/µl)		0.3	0.3	0.3		
JetPEI	6 µl	6 µl	6 µl	6 µl	6 µl	6 µl

The day after, cells were trypsinized using 0.05% trypsin in PBS for 5 minutes at 37°C. Trypsinization is stopped by adding 2mL of DMEM +SVF. Cells are centrifuged at 1200rpm for 5 minutes and cells are resuspended in 1 ml of serum free Opti-MEM medium (ThermoFisher Scientific) to adjust the cell concentration to 600,000 cells per mL. 250µl cells were then mixed together in a 1:1 ratio (400 000 cells/ml), and were agitated at 100 rpm for 2 hours at 37°C.

The below mixes of NF155 and CNTN1/Caspr1 were prepared:

	1	**2**	**3**	**4**
GFP (400,000 cells)	X			
hNF155 (400,000 cells)	X	X		
hNF155-P904R-fs (400,000 cells)			X	
hNF155-R705P (400,000 cells)				X
hCNTN1/Caspr1 (400,000 cells)		X	X	X

100 µl of cell suspension was then mounted between slides and coverslip, and was immediately observed using a fluorescence microscope at 5X objective. The number of cells in mixed aggregates was evaluated in 10 fields. Aggregates were defined as clusters of at least 4 cells including red and green cells. The four experiments were performed for each condition.

Immunoblots

HEK293 cells were transfected with Nfasc155 constructs for 24 hours as described above using JetPEI Polyplus transfection agent. Then cells were washed once in

PBS and were solubilized on ice for 15 min in 1% Triton X100, 140 mM NaCl, 20 mM Tris-HCl pH 7.4, 2 mM EDTA, containing protease inhibitors.

BCA assay is a biochemical assay used for determining the total concentration of protein in a solution, in this case before using them in a blot. A standard curve needs to be produced by using the BCA protein assay kit (Thermo scientific) as per manufacturer's instructions to predict the right protein concentration of each lysate and adjust it to 2.5 mg/ml. 20 μl of sample buffer was added to 20 μl of total protein (50μg) and heated for 1 min at 90°C, followed by the addition of 20 μl of loading buffer (see buffers' recipes in tables 14 and 15).

Table 14. Sample Buffer recipe.

Reagent	**Quantity (ml)**
water	1.5
Lower buffer (0.5 M Tris HCl 6.8)	2.5
SDS 14%	10

Table 15. Loading Buffer recipe.

Reagent	**Quantity (ml)**
glycerol	5
saccharose 50 %	2
bromophenol blue 0.05%	2
ddH20	1

50 μg of proteins were loaded on 7.5% SDS-PAGE gel, firstly running slowly at 80V followed by an increase to 150V. The gel was transferred to a nitrocellulose membrane for 2 hours at 70V. The membrane was blocked with 3-5% milk for 1 hour

at RT, and immunoblotted with a primary mouse antibody against Myc (1:2000) or a mouse antibody against α-tubulin (1:2000; MABT205; Merck) overnight at 4°C. The membranes were then washed 3 times with TBST for 5 mins. Immunoreactivity was revealed using peroxidase-coupled secondary antibodies (1:5,000; Jackson ImmunoResearch), finally washed with TBST, and PBS before being placed in BM chemiluminescence kit reagents (Roche). The integrated densities of each protein band were measured with ImageLab software (Biorad).

Protein Modelling

In collaboration with Dr. Sara Fortuna and Dr. Rita de Zorzi (Trieste, Italy), WT and mutated *NFASC* isoforms were built from structure PDB ID 3P3Y. Each protein was placed in a cubic box and minimized. A water layer of 0.8nm and Na^+ ions to neutralize the system were added, and a second minimization was performed. In all cases we used AMBER99SB-ILDN [73] force field and Simple Point Charge water. On all systems we performed NVT and NPT equilibrations for 100ps, followed by 360ns NPT production run at 300 K. The temperature was controlled with a modified Berendsen thermostat [74], the pressure with an isotropic Parrinello-Rahman at 1 bar. The iteration time step was set to 2fs with the Verlet integrator and LINCS [75] constraint. We used periodic boundary conditions. Configurations were sampled every 10ps. All the simulations and their analysis were run as implemented in the GROMACS package [76].

Immunofluorescence staining of patient skin

3mm skin samples were taken using a disposable punch after intradermal injection of lidocaine from the leg and forearm of patient 1, and from leg, thigh and fingertip of

patient 3, and of a 5-year-old boy with a suspected indifference to pain as control. Specimens were fixed overnight in Zamboni solution (American Mastertech), cryoprotected in 20% sucrose in PBS and sent in a refrigerated package to the laboratory of Telese. The skin samples were cut into 50-µm-thick sections using a freezing slide microtome (Leica 2000R). Free-floating sections were processed for indirect immunofluorescence using antibodies to stain nerve fibres, myelin and vascular structures. To visualize nodal and paranodal architecture, rabbit antibodies against panNeurofascin (Courtesy of Prof P. Brophy; 1:2000), Nfasc186 (Cell Signaling Technology; 1:800) or *CASPR1* (Abcam; 1:1000) were used. Axon and myelin were visualized with primary mouse antibodies against protein gene product 9.5 (PGP; AbD Serotec; 1:800), myelin basic protein (MBP; Santa Cruz Biotechnology; 1:800), or collagen IV (COLIV; Chemicon; 1:800) or a rabbit antibody against PGP (Biogenesis; 1:400). Species-specific secondary antibodies coupled with Cyanine 2 and Cyanine 3 fluorophores were used to visualize the structures of interest. ULEX Europaeus agglutinin 1 coupled with Cyanine 5 was used to visualize blood vessels and epidermis. Skin sections were mounted on coverslips with agar, dehydrated in alcohol, clarified in methylsalicylate and finally mounted in DPX. Digital images were acquired using a non-laser confocal microscope (Apotome; Zeiss).

Co-expression Network Analysis

In collaboration with Dr. Mina Ryten (UCL), co-expression network analysis was used to investigate the function of *NFASC* across the human central and peripheral nervous system. This analysis was performed by using GTEx V6 gene expression data [77; 78] to generate co-expression networks for each of the 13 brain tissues and tibial nerve RNA-seq data included within the GTEx study. The raw FPKM

(Fragments Per Kilobase of transcript per Million mapped reads) values were corrected for known batch effects, age at death, sex and post-mortem interval, as well as unknown effects. The unknown effects were detected with the Surrogate Variable Analysis (SVA) R Package [79] and correction was performed using ComBat [80]. The resulting residuals were used to create a signed network using the blockwiseConsensusModules R function from the WGCNA R package [81] for each of the 14 tissues. Next, the modules obtained in each of the 14 networks were assigned to cell types using the userListEnrichment R function implemented in the WGCNA R package, which measures enrichment between module-assigned genes and defined brain-related lists using a hypergeometric test. The same approach was used to annotate modules with Gene Ontology, REACTOME [82] and KEGG [28] terms.

Methods of Chapter 3.3

Whole exome sequencing (WES)

To investigate the genetic cause of the disease, WES was performed in both the affected female and the two parents (Figure 23A: II-1, I-1, I-2) and analysis was carried out as described before.

Zebrafish handling and maintenance

Zebrafish (*Danio rerio*) were reared in accordance with the Animals (Scientific Procedures) Act 1986. Fish were maintained at 28°C, and embryos were cultured in fish water containing 0.01% methylene blue to prevent fungal growth. Wild-type fish were from the AB strain.

RNA injections

To make RNA transcripts for hSFPQ, hSFPQ^{S660N}, gfp-hSFPQ, and gfp-hSFPQ^{S660N}, DNA sequences were inserted into the plasmid pCS2+ (Addgene). Plasmids were linearized and transcribed using the mMessage mMachine SP6 Transcription Kit (Ambion). Capped RNA was purified using mini Quick spin columns (Roche). Zebrafish embryos were injected at the one-cell stage with 150 pg of hSFPQ, hSFPQ^{S660N} RNA or 130 pg of gfp-hSFPQ, and gfp-hSFPQ^{S660N} RNA.

Whole mount immunohistochemistry

Zebrafish embryos were fixed in 4% paraformaldehyde overnight at 4°C, then washed with PBS+0.8% triton X-100, permeabilized with 0.25% trypsin/PBS, and blocked in 10% goat serum/PBS for one hour. Embryos were incubated in primary antibody at 4°C overnight, washed with PBS, and incubated with secondary antibody at 4°C overnight. Primary antibodies were anti-gfp 1:500(Amsbio) and anti-acetylated tubulin 1:1000 (Sigma). Secondary antibodies were Alexa Fluor 488/568 (Life Technologies). Alexa Fluor 568 phalloidin stain (Life Technologies) was added at 1:1000 during secondary antibody incubation.

In-situ hybridization

In-situ hybridization was performed as previously described[83]. Antisense probes targeting the genes *Otx2 and Gbx2* were created from linearized plasmids using an *in-vitro* transcription reaction according to the manufacturer's instructions using DIG Labelling Mix (Roche) or fluorescein Labelling Mix (Roche). Probes were purified using mini Quick spin columns (Roche) and stored in 5X SSC and 50% formamide. Embryos were fixed overnight at 4°C in 4% PFA and washed in PBS/0.1% Tween-20

(PBST), then dehydrated in methanol and stored at -20°C. For hybridization, embryos were rehydrated in PBST, then treated with 10 ug/ml Proteinase K (Sigma) before further fixation in 4% PFA for 20 minutes. Embryos were then incubated for several hours at 65°C in hybridization buffer: 5xSSC, 1% Tween-20, 0.5 mg/ml torula RNA, 50µg/ml heparin, 0.1% CHAPS, 50% formamide. Embryos were then incubated overnight in probes diluted 1:20 in hybridization buffer. The following day, embryos were washed over several hours at 65°C in first hybridization buffer, then 2xSSC/1%CHAPS, then 0.2%SSC/1%CHAPS, then at room temperature in MAB/0.1%Tween. They were blocked in MAB/0.1%Tween + 2% Blocking Reagent (Roche). Embryos were then incubated overnight at 4°C with 1:4000 anti-digoxigenin (Roche) in blocking solution. After washing for several hours in MAB/0.1% Tween at room temperature, embryos were developed in NBT + BCIP in 0.1M NaCl, 0.1M Tris-Hcl pH 9.5, 0.05M MgCl2, 1% Tween-20. After development, embryos were fixed in 4% PFA for 75 minutes.

Methods of Chapter 4.2

To investigate the genetic cause of the disease, WES was performed in both the affected female and the two parents (Figure 28a: II-1, I-1, I-2) and analysis was carried out as described before.

Methods of Chapter 4.3

Whole exome sequencing

To investigate the genetic cause of the disease, WES was performed in both the affected siblings (Figure 29A, II-1 and II-2) and analysis was carried out as described

before to fit a recessive model (i.e., homozygous or compound heterozygous) and/or identify genes previously associated with neurological phenotypes or postaxial polydactyly [84; 85].

Cell-based RT-PCR

To investigate the functional impact of the identified truncating mutation in DDX59 we performed a reverse-transcriptase polymerase chain reaction (RT-PCR) using RNA samples extracted from lymphoblastoid cell lines derived from patients and an age-matched control (see general molecular methods). Semi-quantitative PCR (semi-qPCR) was performed in 50-µL reaction volume prepared by combining the cDNA template, gene-specific primers (F 5'-GATGTTCCCGTTGATGCTGT-3' and R 5'-GAGCTTTATTCGAGAGCAAAACT-3'), nuclease-free water, and SYBR Green Master Mix. The PCR reaction conditions were: one cycle of 94°C for 4 mins, followed by 37 cycles of 94°C for 45 s, 54°C for 45 s, and 72°C for 50 s. In semi-qPCR experiments, all measurements were made in triplicate, and *GAPDH* was used as an endogenous reference gene, with amplification under the same conditions. The PCR products were then loaded in a 1% agarose gel, and densitometry analysis was carried out. According to in-silico predictions, the homozygous single base deletion (c.185delT; p.Phe62fs*13) in *DDX59* would either lead to an early truncation of the protein or cause non-sense mRNA decay (NMD).

Western blotting

We used western blot assay to measure the SHH protein from patient and control-derived lymphoblastoid cell lines. Protein lysates were obtained from cultured lymphoblasts and total protein concentration was measured by means of the

Bradford assay. Samples were separated on SDS-PAGE using Bis-Tris gradient gels (4–12% NuPAGE, Invitrogen) according to the manufacturer's recommendations, electrophoretically transferred onto Immobilon-P transfer membranes (Millipore). Membranes were immunoblotted with the respective antibody: mouse anti-Shh (sc-365112, Santa Cruz, 1:500) and mouse anti-GAPDH (ab8245, Abcam, 1:5000), at 4°C overnight. Blots were then exposed to horseradish peroxidase-conjugated goat anti-mouse IgG (1706516, Bio-Rad Laboratories, 1:5000) 1 h at room temperature. Blots were developed using ECL-Prime (GE Healthcare), visualized via a ChemiDoc™ Touch Imaging System and analysed using Image Lab 5.2 software (Bio-Rab Laboratories). For the quantifications, measurements were made in triplicate and the signal intensity of the Shh bands was normalized to the signal intensity of GAPDH bands.

Weighted gene co-expression analysis

In collaboration with the Ryten Lab, we created the co-expression network for white matter tissue with a WGCNA[81] optimized pipeline as described elsewhere [86] based on 83 brains from healthy donors and on 19152 Affy Human Exon-v2 probes[87]. This generates a clustering of genes. Module membership of each gene to its module is obtained by correlating its expression with the module eigengene (1st principal component of the module's gene expression). We annotated each module by using gProfiler[88] package with GO (Gene Ontology) REACTOME and KEGG terms. We characterize each module's cell type specificity with a Fisher's exact text for enrichment on specific gene markers. We tested the salmon module in white matter co-expression module for enrichment in Genomes England Panel App (see Web resources) genes with a Fisher's exact test. We test all panels with more than 10

genes of any level of reviewed evidence and remove the Intellectual disability panel. Reported p values are corrected for multiple testing using Bonferroni correction. The bottom-up plot is created with Cytoscape 1.8[89]. The seed genes are added to an empty graph, and then we add the gene most connected to each seed gene, based on TOM (Topology Overlap Measure), until we reach the desired size. The size of the blob reflects its connectivity. The blob's outer colour ring shows the colour of the module the gene belongs to. Seed genes are represented with an inner yellow area are the seed genes and those with blue colour are the context genes. The layout algorithm for disposing the genes in a 2D canvas is Edge-weighted Spring Embedded layout.

Drosophila model

In collaboration with Dr. Mutsuddi (Varanasi, India) in order to better delineate the neurological phenotype of our patients we investigated the role in neuronal development of the loss-of-function mutant mahe, the Drosophila homologue of human *DDX59*. Drosophila stocks used for analysis were w1118 (wild-type), maheEP1347 (hypomorphic allele for mahe), and EPΔmahe d08059 (null mutant for mahe). Embryo collection and immunostaining was done as described previously157. Primary antibodies used were rabbit anti-Mahe, 1:300; rat anti-ELAV, 1:200; and mouse anti-22c10, 1:100. Secondary antibodies used were goat anti-rabbit antibody alexafluor-555, 1:200; goat anti-mouse antibody alexafluor-488, 1:200; and goat anti-rat antibody conjugated with FITC at 1:100 dilution. Immunostained embryos were examined with a Zeiss (Thornwood, NY) LSM 510 Meta laser scanning confocal microscope to visualize neuronal morphology and axonal projections.

Methods of Chapter 4.4

Study participants

Clinical data collection involved a detailed review of medical records, photographs, videos, phone interviews and a clinical re-evaluation by a neurologist. Supplementary Tables 3-5 summarises the clinical and demographic details of the included cases.

Sequencing

Exome sequencing was carried out using a number of methods in different centres with different analysis platforms and pipelines used. Families 5, 6, 8,10 were sequenced at GeneDx, where genomic DNA was extracted from the proband and parents (when available). The exonic regions and flanking splice junctions of the genome were captured using the Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) or the IDT xGen Exome Research Panel v1.0. Massively parallel (NextGen) sequencing was done on an Illumina system with 100bp or greater paired-end reads. Reads were aligned to human genome build GRCh37/UCSC hg19, and analysed for sequence variants using a custom-developed analysis tool. Additional sequencing technology and variant interpretation protocol has been previously described.[90] The general assertion criteria for variant classification are publicly available on the GeneDx ClinVar submission page (http://www.ncbi.nlm.nih.gov/clinvar/submitters/26957/)

Families 17 and 19, Agilent sequence capture was used, as described elsewhere[91]. They were subjected to exome capture with either the Agilent SureSelect Human All Exome 50 Mb kit (Agilent Technologies, Inc., USA) or the Illumina Rapid Capture 37 Mb Enrichment kit. Sequencing with 100-bp paired-end reads was performed using either the Illumina HiSeq2000 or HiSeq4000 instruments (Illumina, Inc., USA), resulting in >94% recovery at 10× coverage and >85% recovery at 20x coverage.

GATK best practices pipeline was used for SNP and INDEL variant identification (http://www.broadinstitute.org/gatk/). Variants were annotated with in-house software1 and homozygous variant prioritization was done using custom Python scripts to keep variants with MAF.

Bioinformatic analysis

Our bioinformatics filtering strategy was carried out as described before to fit either a biallelic or a *de-novo* model to identify variants in genes previously linked to epilepsy, developmental delay, intellectual disability and other neurological disorders. On WES analysis of the index family (F9), the *NARS1* variant c.1633C>T, p.Arg545Cys was picked up according to its frequency and prediction tool scores (SIFT: damaging (score=1), Polyphen: damaging (score=1), GERP: 5.5, Mutation taster: 0.999992). All the candidate variants were further validated as described before by Sanger sequencing.

cDNA and protein sequence variants of *NARS1* are described in accordance with the recommendations of the Human Genome Variation Society using Ensembl ENSG00000134440; ENST00000256854.10 as the reference sequences. Evolutionary conservation of nucleotides was assessed using PhyloP (46 vertebrate species) and genomic evolutionary rate profiling (GERP) scores.[92] These were accessed through the UCSC Genome Browser[93] using genomic coordinates from GRCh37/hg19. Grantham scores were used to assess the physicochemical nature of the amino acid substitutions. *In silico* analyses of sequence variants were performed using the pathogenicity prediction tools SIFT, PolyPhen-2 and Mutation Taster version 2.

Generation of the nrs1 vector

The pJR1-41XU-*nrs1* expression vector was used to express the *Schizosaccharomyces pombe nrs1* gene by amplifying the coding sequence of *nrs1* from *S. pombe* DNA with the Nrs1-PJR-F and Nrs1-PJR-R primers (all primers are provided in Supplementary Table 6) using Phusion HF polymerase from NEB. The PCR product was cloned into XhoI digested pJR1-41XU[94] using CloneEZ from Genscript. Plasmids were sequenced to confirm the correct insertion of the fragment.

Deletion of *nrs1* gene in *S. pombe* cells

JB775 (*h- ade6-M216 ura4-D18 leu1-32*) cells were synchronized, made competent, and transformed as described before.[95] Cells were transformed with the plasmid containing the *nrs1* gene, pJR1-41XU-nrs1, and transformants were selected by growth in EMM + ade + leu, generating the strain MR397. The *nrs1* gene was deleted in MR397 cells with the standard method by homologous recombination with the NatMx6 cassette[96; 97] using the primers Nrs1DelFw and Nrs1DelRv (Supplementary Table 6). Transformants were selected in EMM + Nat with no thiamine to promote the expression of the *nrs1* gene from the plasmid. Deletions were checked by PCR using primers Nrs1ck-L and kanR and Nrs1ck-R and kanF (Table S5). The strain generated was named MR409. MR409 cells were synchronized, made competent, and transformed as described.[95] The plasmids of the pJR-41XL series contained either the empty vector, the *NARS1* wild-type gene, or the variants described. Transformants were selected in EMM + ade strains.

Cell culture

Fibroblasts of affected individuals carrying the homozygous c.50C>T, p.Thr17Met, c.32G>C, p.Arg11Pro, and c.1633C>T, p.Arg545Cys; compound heterozygous c.1067A>C, p.Asp356Ala and c.203dupA, p.Met69Aspfs*4; and corresponding controls were grown in high-glucose Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal bovine serum and 1% penicillin and streptomycin (as described in general methods).

Semi-quantitative RT-PCR for individual lymphoblasts

Total RNA was extracted from immortalized lymphoblasts available from P2 and parents using TRIzol (Zymo research), as per manufacturer's instructions. The concentration and purity of RNA was determined spectrophotometrically. 1 µg of RNA was reverse transcribed to first strand cDNA using random primers and Moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI, USA). GoTaq® Green Master Mix (Promega) was used and PCR reactions were performed with the following protocol: 95°C-2 min (95°C-30 s, 60°C-30 s, 73°C-1 min) for 35 cycles, 73°C-5 min, and 4°C hold. Two exponential curves representing the product formation was determined for both primer pairs. Cycles 28 and 29 were chosen for *NARS1* and *GAPDH* genes respectively so that amplification rates were in the linear range for semi-quantitative comparisons. Reactions were repeated in triplicate.

Western blotting

For western blotting analysis, protein lysates were obtained from cultured fibroblasts and total protein concentration was measured by means of a Bradford assay. Aliquots of total protein (15 µg) were loaded on 4–12% SDS-polyacrylamide gels (NuPAGE 4–12% Bis-

Tris Protein Gels, ThermoFisher Scientific), transferred to polyvinylidene fluoride membranes, and blocked and incubated overnight with a polyclonal antibody recognizing AsnRS1 (anti-rabbit 1:1000; Proteintech). Secondary antibody was added for one hour and signal was detected using ECL reagents (Amersham Biosciences). Anti-beta-actin antibody (Sigma Aldrich, A3853; 1 in 5000) was used as a loading control. Blots were repeated in triplicate and statistics were performed using Prism 6. Data is presented as mean ± standard error of the mean (SD). The significance between the variables was shown based on the p-value obtained (ns indicates $P > 0.05$, * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, **** $P < 0.00005$).

Blue native polyacrylamide gel electrophoresis (BN-PAGE)

Fibroblast pellets were lysed using 10mM Tris (pH 8), 150mM NaCl, 0.1% NP40 with physical agitation for 30 minutes, centrifuged at 8,000xg to remove debris. The supernatant was removed and total protein was quantified with the Bradford assay. Protein concentrations were equalised and prepared to 20µl at 1µg/µl with NativePAGE™ sample buffer (Thermo) and 1ul of NuPAGE 5% G-250 Sample Additive (Thermo), and then loaded to a NativePAGE 3-12% Bis-Tris Protein Gel (Thermo). Proteins were transferred to PVDF membrane using an iBlot2 PVDF Mini transfer stack (ThermoFisher Scientific) and probed with anti-NARS1 monoclonal antibody (Abcam ab129162, 1:5000) and GAPDH (Santa Cruz). Blots were repeated in triplicate and differences analysed by Welch-corrected t-test.

Induced Neuronal Progenitor Cell (iNPC) conversion

Based on the protocol published by Meyer *et al*. [98], iNPCs were generated from primary fibroblasts by transduction with *Oct4*-, *Klf4,* and *c-Myc*-Sendai virus, followed by culturing

in NPC induction media (1:1 DMEM/F-12: Neurobasal, 2x N2, 2x B27, 1% GlutaMAX, 10ng/ml hLIF, 3μM CHIR99021,and 2 μM SB431542). Neuroepithelial colonies were formed after 3-4 weeks of culturing. These were then isolated and expanded before total RNA was extracted from individual fibroblasts, age and sex matched healthy control fibroblasts, and iNPCs cells using the mirVana^{TM} miRNA Isolation Kit (Ambion) for gene expression analysis by qPCR to confirm iNPC lineage and RNAseq in control and individual iNPCs to identify differentially expressed genes.

RNA extractions and qPCR

Cell pellets from individual fibroblasts and iNPCs were lysed by Trizol reagent. Following the addition of chloroform, the aqueous phase was transferred to RNeasy spin column (QIAGEN) for RNA isolation and resuspension. cDNA was generated using the reverse transcriptase kit (Applied Biosystems) and qPCR (Applied Biosystems 7900HT) was performed in triplicates using SYBR Green PCR Master Mix (Invitrogen, 4309155). Samples were normalised to expression of GAPDH and β-actin and repeated in triplicate.

RNA sequencing

Libraries were prepared using Illumina TruSeq Stranded Total RNA with Ribo-Zero Human kit and were sequenced on an Illumina HiSeq 2500 using a paired-end protocol. Quality of sequencing reads were ensured using FastQC. Reads were aligned using STAR aligner and variants were called using the 2-pass protocol outlined in the GATK documentation (see Web resources). The numbers of reads were counted using HTSeq-count[99]. Differentially expressed genes were identified using the DESeq2 Bioconductor package[100]. Differentially expressed genes with a false discovery rate of ≤0.1 and a $\log_2$

(fold change) ≥1 were considered significant. Gene set enrichment analysis was performed using the CPDB web tool.

AsnRS1 enzyme assay

Aminoacylation was assessed by measuring *NARS1* activity in cultured fibroblasts and lymphoblasts. Cell lysates (cytosolic fraction) were incubated in triplicate at 37°C for 10 minutes in a reaction buffer containing 50mmol/L Tris buffer pH 7.5; 12mmol/L $MgCl_2$; 25mmol/L KCl; 1 mg/mL bovine serum albumin; 0.5mmol/L spermine; 1mmol/L ATP; 0.2mmol/L yeast total tRNA; 1 mmol/L dithiothreitol; and 0.3mmol/L [$^{15}N_2$]-asparagine, [$^{13}C_4$,^{15}N]-threonine, [D_2]-glycine, [$^{15}N_2$]-arginine, and [D_4]-lysine. The reaction was terminated using trichloroacetic acid. Ammonia was then added to release the labelled amino acids from the tRNAs. [$^{13}C_2$,^{15}N]-glycine and [$^{13}C_6$]-arginine were added as internal standards and the labelled amino acids were quantified by LC-MS/MS. Intra-assay variation was <15 TARS1, GARS1, RARS1 KARS1 activity were simultaneously detected as control enzymes. AsnRS1 activities were measured blind and testing was repeated in triplicate. Data is presented as mean ± standard error of the mean (SD). The statistical significance of the difference of AsnRS1 activity between controls and individuals/carriers was determined using a Student t-test with a 95% confidence interval using SPSS 26.

Protein modelling

The crystal structure of the *Brugia malayi* AsnRS1 protein with a 65% identity to human *AsnRS1* sequence, stored under the 2XGT code in the Protein Data Bank was used for the molecular modelling analysis. The homology model of the dimeric human AsnRS1 overlapped with the *S. cerevisiae* DARS-tRNAAsp ligase, co-crystallized with tRNA

molecule PDB 4WJ4, thus having a similar domain organization and sharing 27.5% of sequence identity with the human AsnRS1. The protein was visualized with the University of California–San Francisco Chimera software.[101]

Methods of Chapter 4.5

Whole Exome Sequencing

WES analysis was carried out to fit a recessive (homozygous or compound heterozygous) or a *de-novo* model to identify variants in genes previously linked to developmental delay, intellectual disability and other neurological disorders.

Exome capture on subject 2 was carried out using SureSelect Clinical Research Exome V2 (Agilent) and WES raw data were processed and analyzed using an in-house implemented pipeline previously described [102-104]which is based on the GATK Best Practices[105]. The UCSC GRCh37/hg19 version of genome assembly was used as a reference for reads alignment by means of BWA-MEM tool [106] and the subsequent variant calling. High-quality variants were filtered against public databases (dbSNP150 and gnomAD V.2.0.1) so that only clinically associated variants with unknown frequency or having MAF <0.1%, as well as variants occurring with frequency <1% in our population-matched database (~1500 WES), were considered. We used SnpEff v.4.3 [107] and dbNSFP v.3.5 [108] tools for variant functional annotation, including Combined Annotation Dependent Depletion (CADD) v.1.4 [109] (tool for scoring the deleteriousness of single nucleotide variants where a score of >20 indicates the 1% most deleterious), Mendelian Clinically Applicable Pathogenicity (M-CAP) v.1.0[110] (a classifier for rare missense variants where a

sensitivity score of ≤ 0.95 is possibly pathogenic) and Intervar v.2.0.1[111] for functional impact prediction.

TARS1 enzyme assay

In collaboration with Dr. Marisa Mendes (Amsterdam UMC) aminoacylation was assessed by measuring TARS1 activity in cultured fibroblasts and lymphoblasts. Cell lysates (cytosolic fraction) were incubated in triplicate at 37°C for 10 minutes in a reaction buffer containing 50 mmol/L Tris buffer pH 7.5, 12 mmol/L MgCl2, 25 mmol/L KCl, 1 mg/mL bovine serum albumin, 0.5 mmol/L spermine, 1 mmol/L ATP, 0.2 mmol/L yeast total tRNA, 1 mmol/L dithiothreitol, [13C4,15N]-threonine, and [D4]-lysine. The reaction was terminated using trichloroacetic acid (TCA). After sample washing with trichloroacetic acid, ammonia was added to release the labelled amino acids from the tRNAs. [13C2,15N]-glycine and [13C6]-arginine were added as internal standards and the labelled amino acids were quantified by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Glycyl-tRNA synthetase (GARS1), arginyl-tRNA synthetase (RARS1) and lysyl-tRNA synthetase (KARS1) activity were simultaneously detected as control enzymes.

Molecular modelling

With the guidance of Dr. Marie Sissler (University of Bordeaux) homology models were constructed, using PYMOL (see Web Resources). The templates, used as backbone, were the crystal structure of Pyrococcus horikoshii AsnRS (Protein Data Bank access # 1X55) and the dimeric protein form of mt-AsnRS of Entamoeba histolytica (Protein Data Bank access # 3M4Q).

Yeast complementation assay

In collaboration with Rebecca Meyer-Schuman from the Antonellis Lab (University of Michigan) yeast complementation assays to study the functional consequences of *TARS1* patient variants were performed as previously described.[112] The patient variants p.Arg131His, p.Val372Ile, p.Arg663Gln, p.Arg619Cys, and p.Gln639Pro were introduced individually into the *TARS* open reading frame using site-directed mutagenesis; additionally, variants p.Arg131His and p.Val372Ile were modeled in cis, as were p.Arg619Cys and p.Gln639Pro. Mutagenesis was performed in the pDONR221 vector (see Supplementary Table 8). After mutagenesis, the TARS open reading frame was fully sequenced and then recombined into the pYY1 expression construct bearing the LEU2 gene[113] using Gateway cloning (Invitrogen). These constructs were transformed into a haploid ΔTHS1 yeast strain, with viability maintained using an exogenous copy of THS1 on a maintenance vector with the URA3 marker. Transformants were grown on media lacking both leucine and uracil to select for both the maintenance and experimental plasmids. Individual colonies were grown to saturation, shaking at 275 rpm for two days in liquid culture at 30°C. One mL of saturated culture was then centrifuged at 15,000 rpm for 1 minute and resuspended in 50 µl of Ultrapure water. 10 µl of each concentrated culture was spotted undiluted or diluted 1:10, 1:100, or 1:1000 in water on plates containing 0.1% 5-FOA (Teknova) to select for spontaneous loss of the maintenance vector bearing URA3[114]. Plates were incubated at 30°C and yeast growth was visually inspected after 3-5 days. Results for these have not been acquired yet as the experiments are underway and therefore are not included in this thesis.

Methods of Chapter 5.1

To investigate the genetic cause of the disease, trio-based whole exome sequencing (WES) was performed in the family (Figure. 63A: I-1, I-2, II-1) as described before.[115] As part of our filtering strategy, only exonic and donor/acceptor splicing variants were considered. In accordance with the pedigree and phenotype, priority was given to rare variants [<0.01% in public databases, including 1000 Genomes project, NHLBI Exome Variant Server, Complete Genomics 69, and Exome Aggregation Consortium (ExAC v0.2)] that were fitting a recessive (homozygous or compound heterozygous) or a de-novo model and/or variants in genes previously linked to intellectual disability. There were neither homozygous variants nor plausible compound heterozygous variants segregating with the disease in the family. We also excluded other genes' variants that could be related to familiar benign epilepsy and paroxysmal movement disorders (for example *PRRT2*).

Methods of Chapter 5.2

Human participants

Clinical ascertainment included physical examination, medical history interviews, and specialized consultation by paediatric neurologists and clinical geneticists. Two independent neurologists reviewed brain MRIs and EEGs. GPI anchoring studies for Patient 6 was approved by the ethics Institutional Review Board of CHU-Sainte Justine and Université de Montréal.

Exome and Sanger sequencing

WES was performed on genomic DNA extracted from venous blood or buccal samples according to standard procedures and performed as a trio with parental

samples. For all the affected individuals, ES and data analysis was performed as described.[90; 115; 116] For Patient 6, ES was performed in the setting of routine diagnostic testing without the requirement for institutional ethics approval. The candidate variants were confirmed, and segregation analysis was performed by Sanger sequencing (Figure 64A). All variants are listed using transcript NM_033198.4.

Computational and *in vitro* splice analysis

With the guidance of Dr. Barbara Vona (University of Tuebingen) computational analysis of the c.174G>C variant was performed as previously described.[117] To assess the predicted splicing impact of the c.174G>C (p.(Gln58His)) variant in patients 1 and 2, an *in vitro* splicing assay was performed with minor modifications.[118; 119] Briefly, a 1,535 bp genomic region including exons 1 to 3 was PCR amplified from the genomic DNA of a control individual and proband with primers containing *Xho*I and *Bam*HI restriction sites (primers available in Appendix). PCR fragment ligation in a linearized pSPL3-vector preceded transformation into DH5α competent cells (NEB 5-alpha, New England Biolabs) and overnight incubation. The wild-type and mutant-containing vector sequences were confirmed by Sanger sequencing and transfected into HEK 293T cells (ATCC) with FuGENE 6 Transfection Reagent (Promega). Empty vector and transfection negative reactions controls were included. Total RNA was prepared from 24-hour post-transfected cells using a miRNAeasy Mini Kit (Qiagen) that was reverse transcribed using a High Capacity RNA-to-cDNA Kit (Applied Biosystems). Amplified fragments were visualized on a 1% agarose gel (Sigma) and Sanger sequenced. cDNA amplicons

from the wild-type individual were cloned using the TA cloning dual promoter with pCRII kit (Invitrogen).

Flow-cytometry analyses

Fluorescence-activated cell sorting (FACS) analysis to assess for reduction in cell-surface expression of GPI-APs was done for Patient 6 and performed as previously described.[15] This additional work was carried out at Dr. Philippe Campeau's lab (University of Montreal).

Methods of Chapter 5.3

After local institutional review board approval of this study and informed consent from the families, we collected blood samples from the three patients and their parents, and extracted DNA using standard procedures.

Exome sequencing

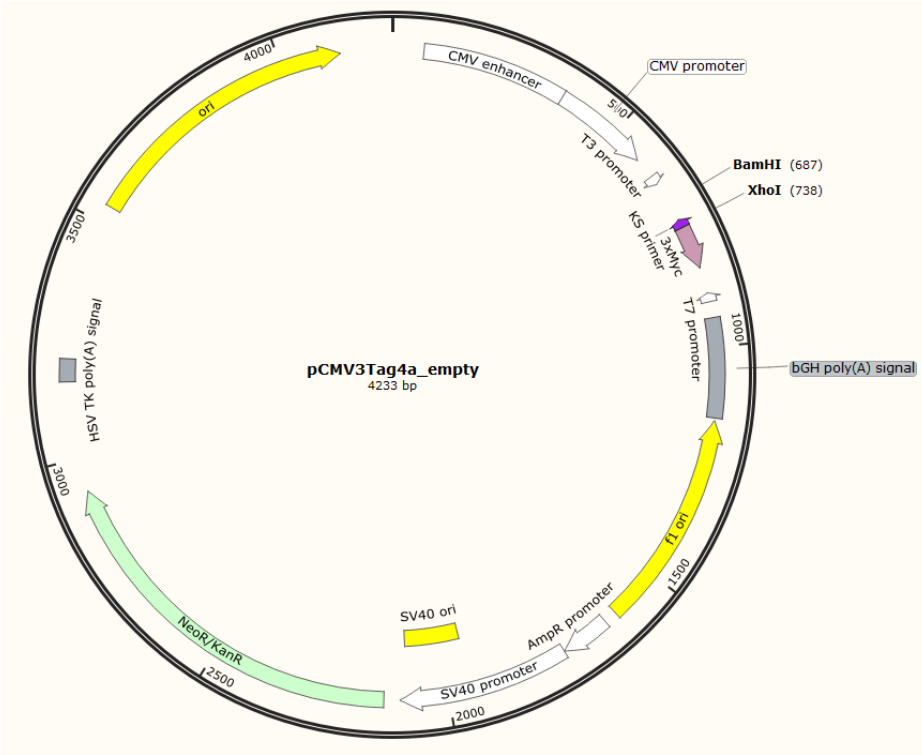
To investigate the genetic cause of the disease, WES was performed in the three affected siblings (Figure 70A, II-1, II-2 and II-3) as described earlier and analysis was carried out to fit a recessive model (i.e., homozygous or compound heterozygous), and/or located in genes previously associated epilepsy or other neurological phenotypes. Genomic DNA from the submitted specimens analysed at GeneDx was enriched for the complete coding regions and splice site junctions for most genes of the human genome using a proprietary capture system developed by GeneDx for NGS with CNV calling (NGS-CNV). Using a custom-developed analysis tool (XomeAnalyzer), data were filtered and analysed to identify sequence variants and most deletions and duplications involving three or more coding exons. [90]

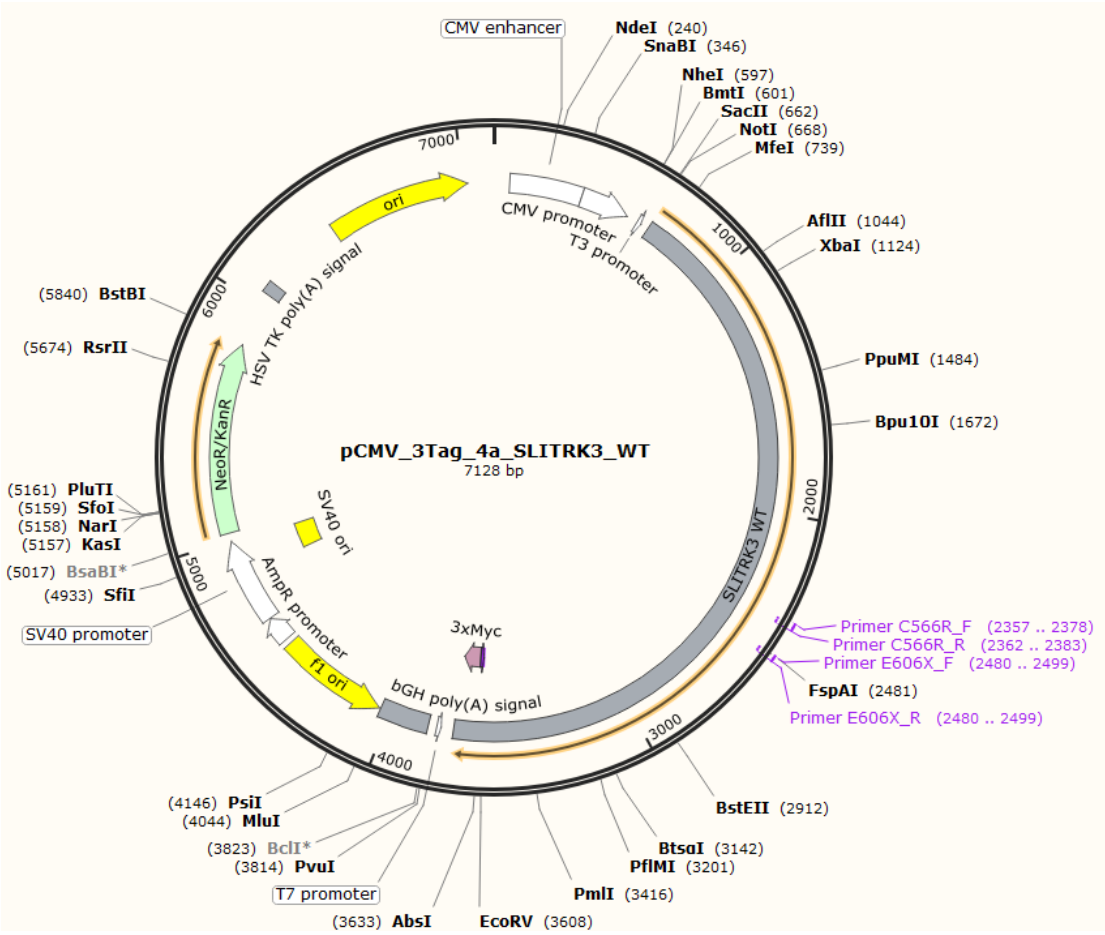
Cloning

ST3 constructs were cDNA synthesised by Genescript (USA) into pCMV-3Tag-4A by replacing the CDs for these inserts via using restriction enzymes BamHI and XhoI (NEB), giving rise to all ST3 plasmid variants (Figure 6).

The variants generated were pCMV_ST3[WT], pCMV_ST3[E606X] and pCMV_ST3[C566R]. Three Myc epitopes were then inserted before the stop codon into the three plasmids with the QuickChange Site-Directed Mutagenesis Kit (Agilent Technologies) as outlined in Chapter 3 with the below primers:

E606X	
Forward	ACTATTtAGCTGGAAGTTCTTTGCC
Reverse	CCAGCTaAATAGTGCGCACATCACG
C566R	
Forward	GAATCCTTGGGACCcGCACCTGT
Reverse	AGGTCACAGGTGCgGTCCCAAG


pCMV3Tag4a_empty
4233 bp
CMV enhancer
CMV promoter
T3 promoter
BamHI (687)
XhoI (738)
KS primer
3xMyc
T7 promoter
bGH poly(A) signal
f1 ori
AmpR promoter
SV40 promoter
SV40 ori
NeoR/KanR
HSV TK poly(A) signal
ori
1000
1500
2000
2500
3000
3500
4000
5000


pCMV_3Tag_4a_SLITRK3_WT
7128 bp
CMV enhancer
NdeI (240)
SnaBI (346)
NheI (597)
BmtI (601)
SacII (662)
NotI (668)
MfeI (739)
AflII (1044)
XbaI (1124)
PpuMI (1484)
Bpu10I (1672)
CMV promoter
T3 promoter
SLITRK3 WT
Primer C566R_F (2357 .. 2378)
Primer C566R_R (2362 .. 2383)
Primer E606X_F (2480 .. 2499)
FspAI (2481)
Primer E606X_R (2480 .. 2499)
BstEII (2912)
BtsαI (3142)
PflMI (3201)
PmlI (3416)
EcoRV (3608)
AbsI (3633)
T7 promoter
PvuI (3814)
BclI* (3823)
MluI (4044)
PsiI (4146)
bGH poly(A) signal
3xMyc
f1 ori
AmpR promoter
SV40 ori
SV40 promoter
SfiI (4933)
BsaBI* (5017)
NarI (5158)
SfoI (5159)
PluTI (5161)
KasI (5157)
NeoR/KanR
HSV TK poly(A) signal
RsrII (5674)
BstBI (5840)
ori
1000
2000
3000
4000
5000
6000
7000

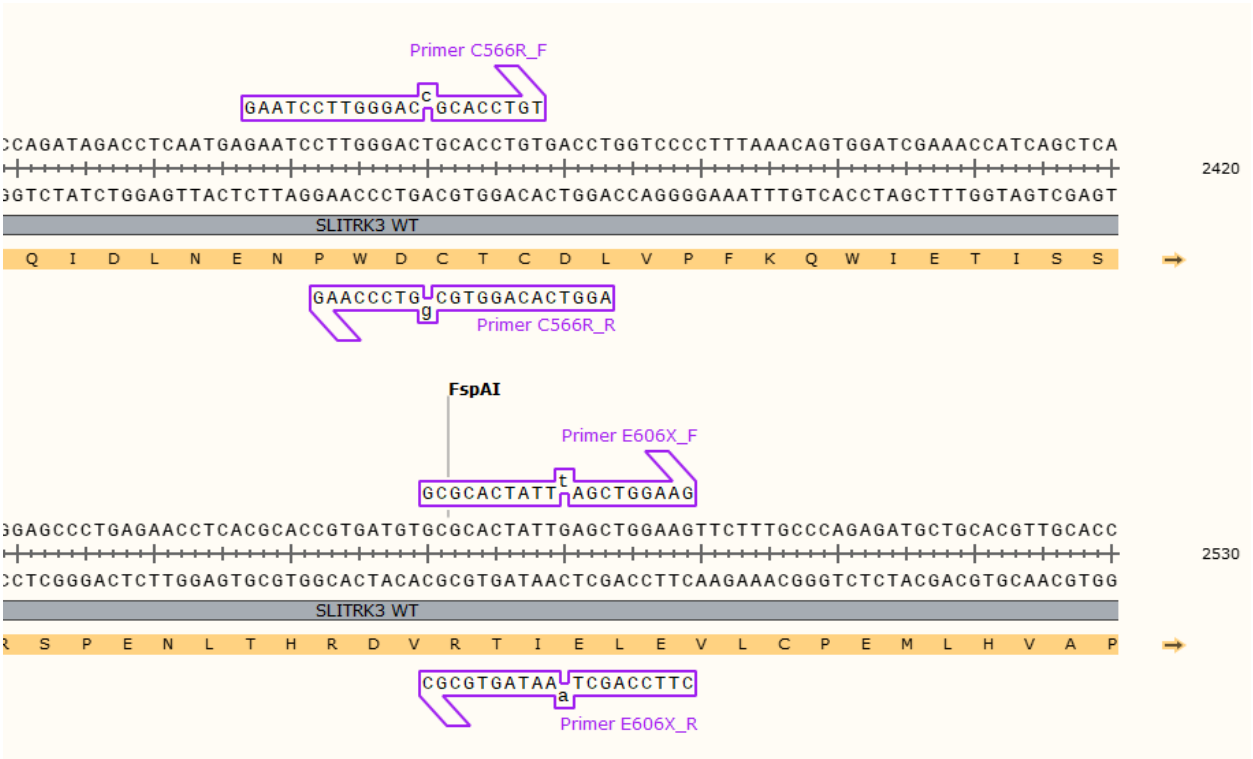


Figure 6. Plasmid maps of *SLITRK3* construct using Snapgene software. An empty pCMV-3Tag-4A plasmid (top) was used as a template for all SLITRK3 isotope plasmids. An example plasmid with a *SLITRK3* wild-type (top) as generated by Genscript (USA) using restriction enzymes BamHI and XhoI (NEB) to replace the CDs for the WT insert and create the ST3_E606X and ST3_C566R plasmids by using the mentioned mutagenesis primers.

Western blot

Protein concentration was quantified with the standard BCA method. An equal amount of loading samples was mixed with an equal volume of 2x SDS loading buffer and boiled for 5 min at 95°C, and then separated using pre-casted 10% SDS-PAGE gels (BioRad). The proteins were transferred onto PVDF membranes, blocked, and incubated with primary antibodies (rabbit polyclonal anti-SLITRK3, Proteintech, Cat#: 21649-1-AP and rabbit monoclonal anti-Myc-Tag (71D10), Cell

Signaling Technology, Cat#: 2278,) overnight at 4°C. The PVDF membranes were then washed three times with 0.1% TBST and incubated with HRP-conjugated secondary antibodies for 1 hr at RT. Protein was detected with the standard enhanced chemiluminescence (ECL) method.

Primary Hippocampal Neuron culture

In collaboration with Dr. Wei Lu's lab (NIH, Bethesda) primary cultures of hippocampal neurons were prepared as described previously. [120] Briefly, timed-pregnant mice at E17.5-18.5 were anesthetized on ice and decapitated. The fetal hippocampi were quickly dissected out in ice-cold Hank's balanced salt solution, triturated with a sterile tweezer, and digested with papain (Worthington, LK003176) solution at 37°C for 30 min. After 5 min centrifuging at 800 r.p.m. at room temperature (RT), the pellet was resuspended in DNaseI containing Hank's solution and then was mechanically dissociated into single cells by gentle pipetting up and down. Cells were then transferred into Hank's solution mixed with trypsin inhibitor (10 mg/ml, Sigma, T9253) and BSA (10 mg/ml, Sigma, A9647), and centrifuged at 800 r.p.m. for 10 min. The pellet was resuspended in neurobasal plating media with 2% fetal bovine serum (FBS) (Gibco, 10437-028), 2% B27 supplements, and L-glutamine (2 mM). Neurons were plated at a density of ~0.9 x 105 cells/well on poly-D-lysine (Sigma, P6407) pre-coated glass coverslips (12 mm) residing in 24-well plates for electrophysiological recording and immunocytochemistry. Culture media were changed by half volume with neurobasal maintenance media containing 2% B27 (GIBCO, 17504-044) supplements and L-glutamine (2 mM) once a week.

Electrophysiology in neurons

For mIPSC recording, ST3-WT, ST3-E606X, or ST3-C566R together with pCAGGS-IRES-GFP (ST3:GFP=9:1) were transfected into cultured hippocampal neurons at DIV 14-15 using Lipofectamine 3000 (ThermoFisher, L3000015). 48 hours after transfection, coverslips were transferred to a submersion chamber on an upright Olympus microscope, and perfused with aCSF solution supplemented with TTX (0.5 μM), DNQX (20 μM) in an external solution (in mM): 140 NaCl, 5 KCl, 2 CaCl2, 1 MgCl2, 10 HEPES and 10 D-glucose, adjusted to pH 7.4 with NaOH with 5% CO2. GFP fluorescent positive and negative neurons were identified by epifluorescence microscopy. Neurons were voltage-clamped at −70 mV for the detection of mIPSC events. The internal solution for mIPSC recording (in mM): CsMeSO4 70, CsCl 70, NaCl 8, EGTA 0.3, HEPES 20, MgATP 4, and Na3GTP 0.3. Osmolality was adjusted to 290-295 mOsm and pH was buffered at 7.25–7.35. Series resistance was monitored and not compensated, and cells in which series resistance varied by 25% during a recording session were discarded. Synaptic responses were collected with a Multiclamp 700B amplifier (Axon Instruments, Foster City, CA, United States), filtered at 2 kHz, and digitized at 10 kHz. All recordings were performed at RT. A total of 100–300 consecutive miniature events were semi-automatically detected by off-line analysis using customized software Igor Pro (Wavemetrics) using a threshold of 6 pA. All mIPSC events were visually inspected to ensure that they were mIPSCs during analysis, and non-mIPSC traces were discarded. All pharmacological reagents were purchased from Abcam, and other chemicals were purchased from Sigma.

Chapter 3. Demyelinating Neuropathies

3.1 Clinical and molecular aspects of neuropathies

Inherited neuropathies are clinically and genetically heterogeneous disorders. They mainly affect the peripheral nervous system, and they are characterised by progressive distal muscle atrophy, often with foot deformity, weakness and sensory loss. Inherited peripheral neuropathies can occur as a "pure" neuropathy or as part of a more complex neurological phenotype. Charcot Marie Tooth disease (CMT) and the related neuropathies, distal hereditary motor neuropathy (dHMN) and hereditary sensory neuropathy (HSN) are the classical "pure" neuropathies. [14; 121; 122] They are the commonest group of inherited neuromuscular diseases, and they exhibit wide phenotypic and genetic heterogeneity. In many conditions, inherited neuropathy is part of a multisystem disorder affecting also the CNS or other non-neurological systems.

In the past decade, next-generation sequencing (NGS) technologies have revolutionized genomic medicine and, as these technologies are being applied to clinical practice, they are changing our diagnostic approach to inherited neuropathies. WES has served as an efficient strategy to identify novel genes associated with rare Mendelian disorders. [123] [51] WES is a tool that could be employed diagnostically because of the genetically heterogeneity of the disease. It was first successfully used in the identification of a novel *GJB1* mutation in a CMT family [124]. Several studies subsequently documented WES as a powerful tool in the identification of novel CMT genes [125] and interestingly in the past two decades, diagnostic laboratories and clinicians have together managed to diagnose around 200 genes for a neuropathy as part of a complex inherited syndrome and almost 100

additional genes cause a form of CMT (Figure 7). [126] The application of NGS technologies in the diagnosis of CMT, can also include disease-specific gene panels, whole-genome sequencing, mitochondrial sequencing and high-throughput transcriptome sequencing.

Remarkably the list of genes that need to be screened will only grow bigger as the clinical phenotypes and genetic aetiologies overlap greatly between these disorders. Nevertheless, a definitive genetic diagnosis remains unsolved for many rare cases suggesting the high heterogeneity of inherited neuromuscular disorders.

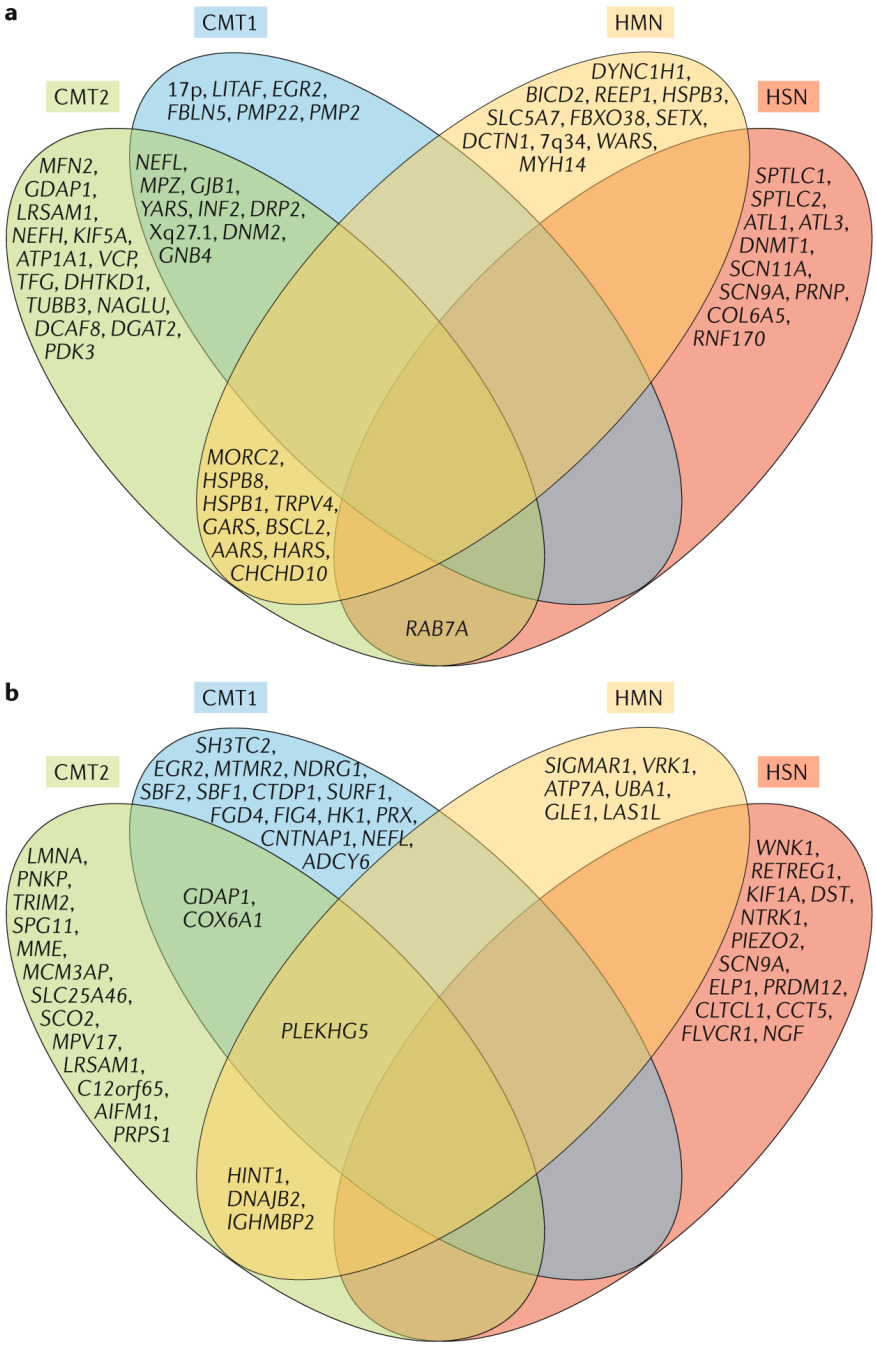


Figure 7. CMT-associated genes by broad CMT phenotype. **a**) Autosomal and X-linked dominant genes. **b)** Autosomal and X-linked recessive genes. Genes linked with each CMT phenotype are contained within the ovals of the corresponding colour, while those which been linked to more than one phenotype are shown in the overlap areas (Pipis *et al.*, 2019).

3.2 Biallelic mutations in neurofascin cause neurodevelopmental impairment and peripheral demyelination

3.2.1 Introduction

In the vertebrate nervous system, myelinated fibres allow rapid nerve impulse transmission because of the ensheathment of their axons by specialized glial cells to form the multilamellar myelin sheath. Also, voltage-gated sodium channels (Nav) are clustered at the nodes of Ranvier further facilitating the rapid conduction of nerve impulse in the vertebrate brain and peripheral nerves. [127] Neurofascin protein isoforms Nfasc186 and Nfasc155 are splice variants encoded by *NFASC* gene (MIM: 609145) and belong to the L1 family of immunoglobulin cell adhesion molecules, which play a critical role in the assembly of the node of Ranvier. [127]

Nfasc155 is a glial isoform abundantly expressed in myelinating Schwann cells and predominantly enriched at the paranodal sites, acting as a ligand for the axonal Contactin-associated protein-1 (CASPR1)/contactin-1 (CNTN1) complex. Nfasc186 represents the axonal isoform expressed at the node of Ranvier membranes where it is involved in the long-term maintenance and stability of Nav channels (Figure 8). [128; 129] Autoantibodies against Nfasc186 and Nfasc155 are associated with aggressive forms of chronic inflammatory demyelinating polyneuropathy (CIDP) and cause an immune-mediated nodo-paranodopathy [130; 131] further demonstrating that these proteins play crucial role in conduction. Biallelic variants in gliomedin (*GLDN*) and *CASPR1* encoding essential components of the nodes of Ranvier, and paranodes, respectively, lead to inherited nodo-paranodopathy, a distinct disease entity among peripheral neuropathies. [132] Smigiel *et al.* (2018) report the identification of a

homozygous truncating *NFASC* mutation affecting the fibronectin type III domain, specific to the Nfasc155 isotope in a child presented with a very severe neurodevelopmental disorder resembling spinal muscular atrophy while Monfrini *et al.* (2019) report a homozygous missense mutation in a case with autosomal recessive ataxia and a demyelinating neuropathy. [2; 4] Importantly, a human Mendelian disease caused by *NFASC* variants has been reported only in these 2 single individuals so far and there is not yet a confirmed disease association for this gene in the OMIM database (*NFASC*; MIM: 609145).

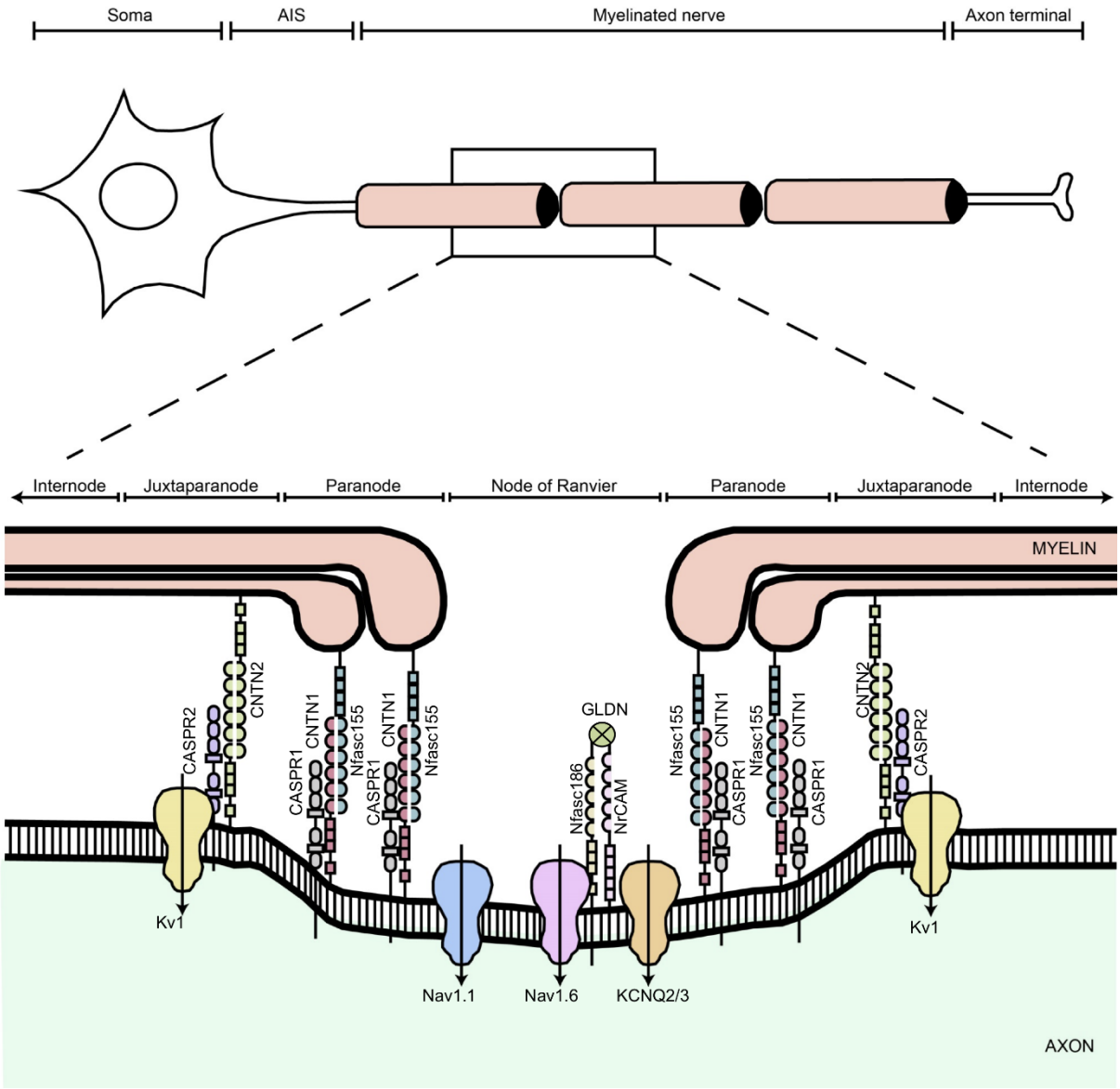


Figure 8. Schematic diagram showing the different domains of a myelinated axon. The axonal region around the node of Ranvier is expanded to show the different axonal domains: the node of Ranvier where voltage-gated Na+ channels (Nav1.6 and Nav1.1) are expressed, the paranode where the myelin is attached to the axon, and the juxtaparanode where most voltage-gated K+ channels (KCNQ2/3 and Kv1) are located. Each of these domains is characterised by the expression of specific cell adhesion molecules; at the nodes Nfasc186 binds Gliomedin (GLDN) and NrCAM which are secreted by Schwann cells in the nodal gap lumen, at the paranode Nfasc155 forms a complex with CNTN1/CASPR1 to forms the septate-like junctions, and at the juxtaparanode the CNTN2/CASPR2 complex enables the sequestration of Kv1 channels (adapted from Arancibia-Carcamo and Attwell, 2014).

Here, we describe ten individuals from 6 families carrying homozygous non-synonymous or frameshift variants in *NFASC.* Phenotypic analysis of affected individuals revealed in all cases global developmental delay and weakness with variable features of chronic demyelinating neuropathy in some cases. For some of the identified variants, we compared surface expression of the disease variant and wild-type (WT) *NFASC* 155 isoforms. We also tested the capability of variant proteins to interact with other partner proteins at paranodal sites, including the *CNTN1/CASPR1* complex. An absence of paranodal junction was identified by immunofluorescence analyses of myelinated fibres from skin samples of an affected individual carrying a homozygous missense variant in *NFASC*. Our results link biallelic variants in *NFASC* isoforms at glial cells to defects in the paranodal axoglial junctions and phenotypes which range from variable neurodevelopmental impairment to weakness, central hypomyelination, and peripheral chronic

demyelinating neuropathy.

With the help of GeneMatcher, we identified 10 individuals from 6 unrelated families, exhibiting a neurodevelopmental disorder characterized with a spectrum of central (intellectual disability, developmental delay, motor impairment, speech difficulties) and peripheral (early onset demyelinating neuropathy) involvement, who were found by exome or genome sequencing to carry one frameshift (P939Ter) or four different homozygous non-synonymous variants (P694T, N130D, S820P, R359P, V1077A) in *NFASC*.

In collaboration with Dr. Jerome Devaux (Montpelier, France), expression studies using immunostaining-based techniques identified absent expression of the Nfasc155 isoform as a consequence of the frameshift variant (P939Ter) and a significant reduction of expression was also observed in association with two non-synonymous variants (P705T, S831P) affecting the fibronectin type III domain. Cell aggregation studies revealed a severely impaired Nfasc155-*CNTN1/CASPR1* complex interaction as a result of the identified variants. In collaboration with Dr. Maria Nolano (Naples, Italy), immunofluorescence staining of myelinated fibres from two affected individuals showed a severe loss of myelinated fibres and abnormalities in the paranodal junction morphology. Our results establish that recessive variants affecting the Nfasc155 isoform can affect the formation of paranodal axoglial junctions at the nodes of Ranvier. The genetic disease caused by biallelic *NFASC* variants includes neurodevelopmental impairment and a spectrum of central and peripheral demyelination as part of its core clinical phenotype. These findings support possible overlapping molecular mechanisms of paranodal damage at

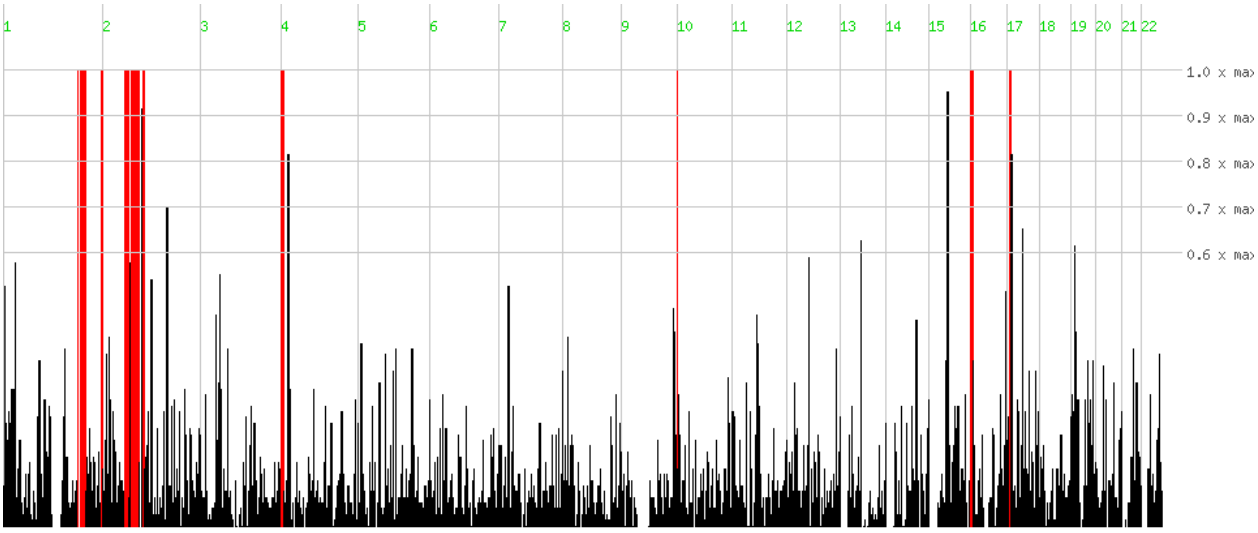
peripheral nerves in both the immune-mediated and the genetic disease, but the observation of prominent central neurological involvement in *NFASC* biallelic variant carriers highlights the importance of this gene in human brain development and function.[1; 6]

3.2.2 Results

Genetic analyses, variant calling, filtering and interpretation

In the index family (Family 1) trio-WES was performed and an average of 20 variants were identified per individual. As this was a family showing a recessive model of inheritance (due to the high degree of consanguinity in the parents) homozygosity analysis was performed from trio-WES data using web-based tools (www.homozygositymapper.org) and identified different regions of homozygosity in the proband in chromosome 1 (186919863-200312546; 200550243-205272722 ; 245530282-248604452), chromosome 2 (52929780-65200618; 69688727-72742083; 73280303-89890648), chromosome 4 (4239539-7312145), chromosome 9 (139440852-139973629), chromosome 16 (1306986-3077710; 3119304-7703785) and chromosome 17 (8007650-10555061) Rare variants with a frequency <0.01% in public databases (including 1,000 Genomes project and ExAC) present as homozygous in the proband and heterozygous in the unaffected parents were selected. This restricted our list candidate variants to homozygous variants in *LAD1* (MIM: 602314; NM_005558.3: c.853A>G; p.Thr285Ala), *ZBED6* (MIM: 613512; encoding a transcriptional repressor containing a zinc-finger domain, *USP43* (MIM: 615695; NM_014709.3: c.4706C>T; p.Ala1569Val) and *NFASC* (NM_001005388: c.2080C>A; p.P694T). The homozygous variant identified in *NFASC* affects an highly conserved residue within the fibronectin (type III) domain and predicted to be

deleterious (according to SIFT and Polyphen tools seen in Table 16). It was prioritized based on the biological importance of this gene in the nodal and paranodal organization and the knowledge of demyelinating neuropathy phenotypes (similar to our patient) in patients with IgG4 antibodies against Nfasc155 and Nfasc186 isoforms (Figure 9). Sanger sequencing segregation analysis confirmed segregation and identified the variant as heterozygous in the unaffected sister.



80	1	186919863	200312546	n/a	n/a	region	genotypes
80	1	200550243	205272722	n/a	n/a	region	genotypes
80	1	245530282	248604452	n/a	n/a	region	genotypes
80	2	52929780	65200618	n/a	n/a	region	genotypes
80	2	69688727	72742083	n/a	n/a	region	genotypes
80	2	73280303	89890648	n/a	n/a	region	genotypes
80	2	99779642	103334969	n/a	n/a	region	genotypes
80	4	67857	3589623	n/a	n/a	region	genotypes

80	4	4239539	7312145	n/a	n/a	region	genotypes
80	9	139440852	139973629	n/a	n/a	region	genotypes
80	16	1306986	3077710	n/a	n/a	region	genotypes
80	16	3119304	7703785	n/a	n/a	region	genotypes
80	17	8007650	10555061	n/a	n/a	region	genotypes

Figure 9. Shared regions of homozygosity were identified using Homozygosity Mapper. Homozygosity mapping in patient 1 revealed several homozygous regions and the region (highlighted in yellow) on chromosome 1 includes NFASC.

Identification of the NFASC variants

Biallelic *NFASC* variants were independently identified as the primary candidate genetic cause for the phenotype observed in the 6 families described herein (Figure 10, Table 16). WES carried in Family 1 revealed a homozygous variant in *NFASC* (chr1:204948591C>A). This variant is predicted to cause a non-synonymous substitution in the two major *NFASC* transcripts: NM_001160331.1: c.2113C>A (p.P705T) and NM_001005388.2: c.2080C>A (p.P694T). NM_001005388.2 is the canonical *NFASC* transcript encoding the longest protein isoform of 1241 amino acids corresponding to Nfasc186, while NM_001160331.1 encodes the protein isoform of 1165 amino acids corresponding to Nfasc155. Sanger sequencing-based segregation analysis was performed in Family 1 confirming the variant. Through Gene Matcher and our network of collaborators, we identified 7 additional cases, with clinical phenotypes partially overlapping that of Family 1. WGS was carried out in proband 2 (Figure 10) at Genedx, Gaithersburg, USA, by massively parallel

(NextGen) sequencing on an Illumina sequencing system with 100 bp or greater paired-end reads. This revealed a homozygous single base deletion (chr1:204986105delC) predicted to cause a frameshift variant in the Nfasc155 isoform NM_001160331.1: c.2816delC (p.P939Ter) and in *NFASC* transcript 3 NM_015090.3: c.2771delC (p.P924Ter). The frameshift variant is expected to cause loss of normal protein function either through protein truncation or nonsense-mediated mRNA decay. WES performed for one individual of Family 3 (Figure 10A) detected a homozygous variant (chr1:204923488A>G), resulting in a missense change in both major transcripts: NM_001005388.2: c.388A>G (p.N130D) and NM_001160331.1: c.298A>G (p.N124D), located within a 10 MB stretch of homozygosity. WES in the proband of Family 4 (Figure 10) revealed a homozygous variant (chr1:204951136 T>C) resulting in a missense change in both NM_001005388.2: c.2458T>C (p.S820P) and NM_001160331.1:c. 2491T>C (p.S831P). The index case of Family 5 (Figure 10) was subjected to WES as described elsewhere [3]. A homozygous missense variant in *NFASC* was identified in chr1:204939816G>C. This variant is predicted in both major transcripts as a non-synonymous substitution: NM_001160331.1: c.1109G>C (p.R370P) and NM_001005388.2: c.1076G>C (p.R359P) and it segregates with the disease in the family (Patients 7 and 8). Family 6 was identified from the replication cohort screening, where WES carried out in probands 9 and 10 and their parents revealed a homozygous missense variant in chr1:204971817T>C. This variant is predicted a non-synonymous substitution only in the NF186 transcript: *NFASC* (NM_001005388.2:c.3230T>C (p.V1077A) and it segregates with the disease in the family.

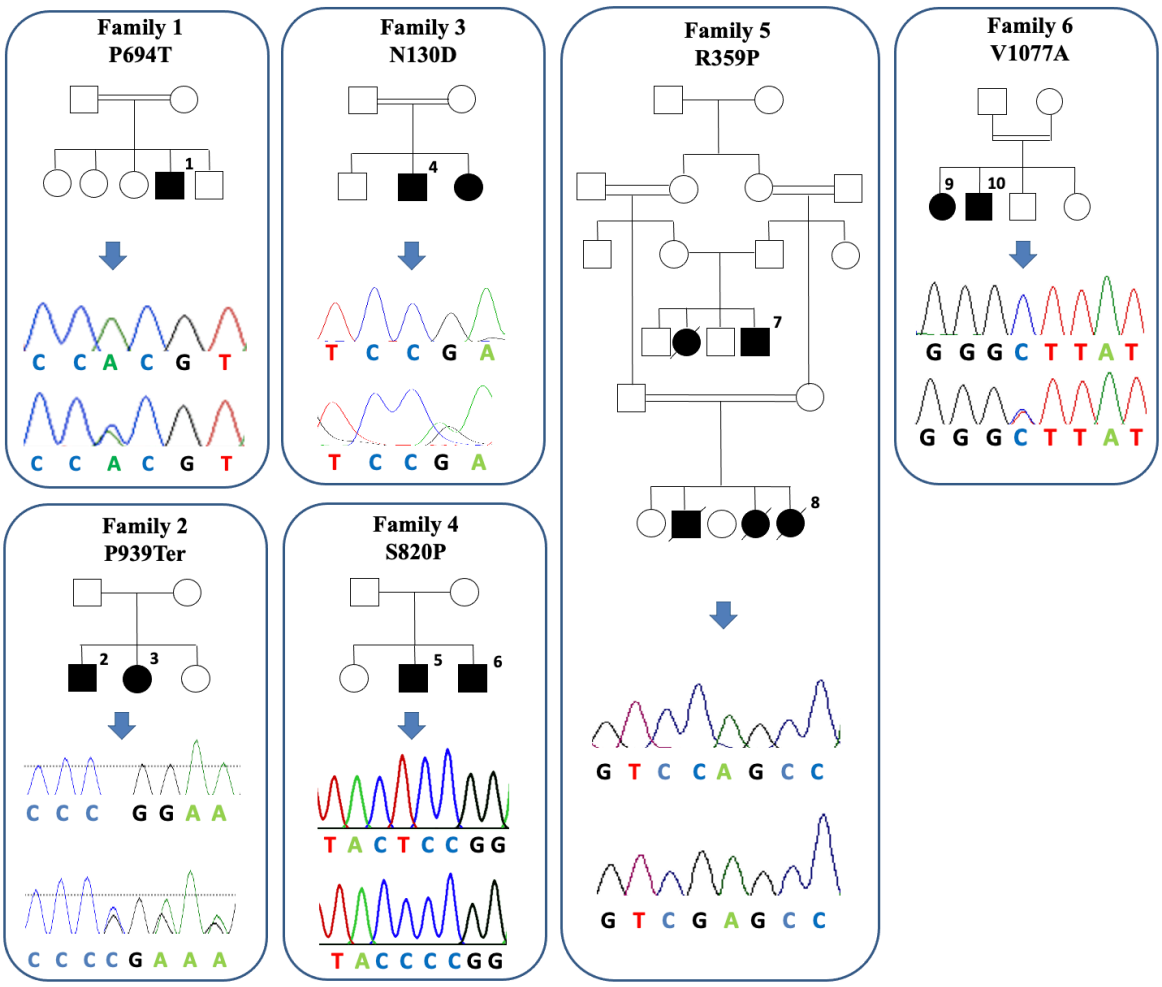


Figure 10. Pedigrees and Sanger sequencing of *NFASC* variants. Pedigrees of the 6 families carrying biallelic *NFASC* mutations and Sanger sequencing electropherograms confirming the mutations (modified from Efthymiou *et al*., 2019).

Family	Genomic coordinates (GRCh37/hg19)	Variant	dbSNP 138	1000G	ESP6500	ExAC	gnomAD	SIFT	PolyPhen	Condel	CADD_PHRED	GERP++
1	1:204948591-204948591	ENST00000339876.6 Nfasc186:c.2080C>A:p.P694T Nfasc155:c.2113C>A:p.P705T	-	-	-	-	-	Deleterious (0)	Probably damaging (0.998)	Deleterious (0.919)	25.4	5.24
2	1:204986105-204986106	ENST00000430393.5 Nfasc155:c.2816delC:p.P939Ter Nfasc3:c.2771delC:p.P924Ter	-	-	-	-	-	-	-	-	-	-
3	1:204923488-204923488	ENST00000339876.6 Nfasc186:c.388A>G:p.N130D Nfasc155:c.298A>G:p.N124D	-	-	-	-	-	Tolerated (0.07)	Possibly damaging (0.67)	Deleterious (0.556)	21.5	5.37
4	1:204951136-204951136	ENST00000339876.6 Nfasc186:c.2458T>C:p.S820P Nfasc155:c.2491T>C:p.S831P	-	-	-	-	-	Deleterious (0)	Probably damaging (1)	Deleterious (0.945)	28.6	5.55
5	1:204939816-204939816	ENST00000339876.6 Nfasc186:c.1076G>C:p.R359P Nfasc155:c.1109G>C:p.R370P	-	-	-	-	-	-	Probably damaging (0.812)	Deleterious (0.75)	34	5.64
6	1:204971817-204971817	ENST00000339876.6 Nfasc186:c.3230T>C:p.V1077A	-	-	-	-	-	Tolerated (0.09)	Probably damaging (0.986)	Deleterious (0.73)	22.7	5.63

Table 16. *NFASC* intragenic variants identified in our cohort. A CADD score ≥ 20 indicates that the variant is predicted to be the among the 1% most deleterious substitutions in the protein-coding parts of the human genome. A GERP++ score of close to 6 indicates a high evolutionary conservation of the *NFASC* sequence across species.

Presentation of the Neurofascin disease

Table 17 summarizes the core phenotypic features of 10 individuals (aged between 3 months and 21 years), including two patients from a Saudi Arabian family (Patients 7 and 8) previously described [3]. Figures 8B and 8C illustrates the clinical features and MRI scans of affected individuals from each family. In all patients, psychomotor development was severely delayed (except patients 5 and 6) in all domains. Affected individuals did not communicate, and none achieved purposeful hand movements or independent ambulation. Hypotonia was present since neonatal age in all cases. Impaired social interaction and only brief and occasional visual contact were noticed in all affected individuals by the first year of age. Several patients developed restricted patterns of interests and repetitive behaviors and frequently exhibited hand or head stereotypic movements. Severe intellectual disability was documented in all cases. None of the affected individuals attained intelligible speech.

Due to muscle weakness and absent reflexes, individuals from families 1, 2, 4 and 6 underwent detailed neurophysiological investigations as part of their diagnostic work up and these showed a severe reduction in peripheral nerve conduction velocities in all cases. Patients from Families 1 and 2 exhibited the most profound phenotype with severe demyelinating and axonal neuropathy at neurophysiological investigations. Nonspecific EEG abnormalities including slow background activities were observed in several families although affected individuals had no history of seizures. *NFASC* is expressed in various brain regions, with the highest transcript level in the white matter region (Figure 1).

Patient	1	2	3	4	5	6	7	8	9	
Gender	M	M	F	M	M	M	M	F	F	
Origin	Moroccan	American	American	Arab Iraqi	Israeli	Israeli	Saudi Arabian	Saudi Arabian	Algerian	
Current age	4 years	1 year and 6 months	10 years	3 years	12 years	16 years	2 years 6 months	4 months (at time of death)	21 years	
Variant	c.2080C>A; p.P694T	c.2816delC; p.P939Ter	c.2816delC; p.P939Ter	c.388A>G; p.N130D	c.2458T>C; p.S820P	c.2458T>C; p.S820P	c.1076G>C; p.R359P	c.1076G>C; p.R359P	c.3230T>C:p. V1077A	c.
Age at onset of initial symptoms	neonatal hypotonia	neonatal hypotonia	neonatal hypotonia	neonatal hypotonia	hypotonia during the first 6 months of life	hypotonia during the first 2 months of life	hypotonia during first 3 months of life	hypotonia during first 3 months of life	neonatal hypotonia	hy xi 1
Developmental Delay	+	+	+	+	+	+	+	+	+	
Intellectual Disability	+	+	+	+	+	+	+	+	+	
Speech impairment	+	+	+	+	+	+	+	+	+	
Hypotonia / weakness	+	+	+	+	+	+	+	+	+	
Neurological examination	weakness, absent reflex, abnormal sensation, no walk	weakness, absent reflex, no walk	weakness, no walk	weakness, no walk	weakness, no walk	weakness, no walk	weakness, no walk	weakness, no walk	no walk	wa wa
Sensory NCS	reduced CV, normal Amp	NT	NT	NT	SSR in hand and food	NT	Normal	NT	normal CV, reduced Amp in ulnar nerve	
Motor NCS	reduced CV (except normal CV ulnar nerve), normal Amp (except peroneal nerve reduced)	NT	severely reduced CV, normal Amp (except peroneal nerve reduced)	NT	NT	NT	Normal	NT	Normal	
EMG	NT	NT	acute and chronic denervation, PSW	NT	NT	NT	Normal	NT	chronic denervation-reinnervation	d re
MRI	atrophic changes and white matter loss	cerebral white matter loss	cerebral white matter loss, severe atrophy of corpus callosum	NT	normal	normal	diffuse white matter T2 hyperintensity	NT	cerebellar atrophy	

Table 17. Genetic, clinical and neurophysiology details of *NFASC* patients. **Genetic, clinical and neurophysiology de**
Abbreviations are as follows **+**, present; **−**, absent; **NCS**: nerve conduction studies. **CV**: conduction velocity/velocities, **Amp**: applicable, **PSW**: positive sharp waves, **SSR**: sympathetic skin response.

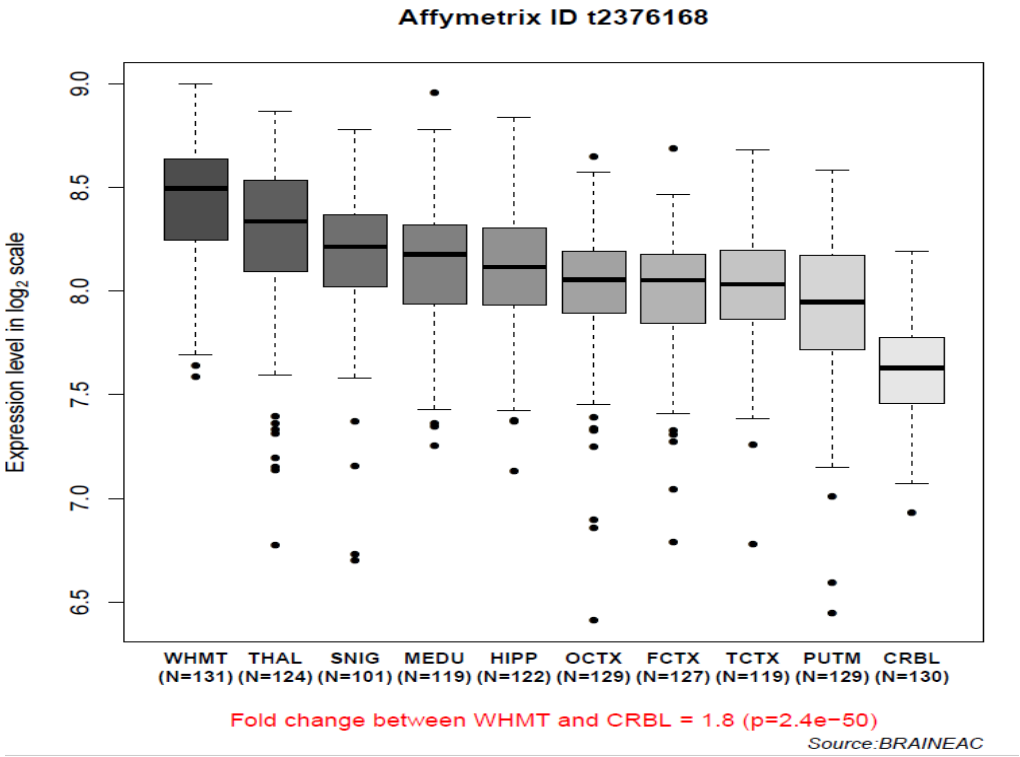


Figure 11. NFASC expression in various brain regions, with the highest transcript level in white matter (from the BRAINEAC). WHMT, intralobular white matter; THAL, thalamus (at the level of the lateral geniculate nucleus); SNIG, substantia nigra; MEDU, inferior olivary nucleus (sub-dissected from the medulla); HIPP, hippocampus; OCTX, occipital cortex; FCTX, frontal cortex; TCTX, temporal cortex; PUTM, putamen (at the level of the anterior commissure); CRBL, cerebellar cortex (Efthymiou *et al.*, 2019).

MRI studies of patients 1, 2, 3, 7, 9 and 10 showed loss of cerebral white matter (periventricular white matter hyperintensities on T2 MRI images) which was more severe in some affected individuals (Patients 1, 2 and 3) and also associated with atrophic changes mainly involving the corpus callosum and the brainstem (Figure 9C). T2-weighted coronal MRI showing enlarged lateral ventricles in Patient 1 indicative of cortical volume loss (Figure 9C; left panel). Similar MRI patterns are also present in *CASPR1*-deficient patients with brain stem and corpus callosum atrophy. [133] Basket cell function is well characterized;

basket cell axons initially form synaptic contacts on the somata of Purkinje cells (PCs). Then, they appear to move directly to the axon initial segment (AIS). Upon reaching the AIS, basket cell axons extend multiple terminal branches and establish the pinceau. Cerebellar atrophy has been previously directly associated to loss of *NFASC* in both Purkinje and basket neurons causing abnormal basket axon collateral branching and targeting to AIS, leading to extensive pinceau disorganization, Purkinje neuron degeneration and severe ataxia. [134] Because of the weakness of the facial muscles, the majority of patients had non-expressive/hypomimic face, and several craniofacial dysmorphisms including bitemporal narrowing, high and wide nasal bridge, micrognathia, glossoptosis and highly arched palate (Figure 9B). The two American children (P939Ter) had more severe phenotype and also developed profound sensorineural hearing loss, feeding difficulties requiring gastrostomy and respiratory difficulties. The younger of the two Saudi Arabian children exhibited skeletal chest deformities and generalized hypertonia (mainly affecting upper limbs) and died at the age of 4 months, presumably due to aspiration defects. The two French children identified from replication cohort screening have an older age of onset compared to the rest of the patients, and the fact that the NFASC transcript NF155 is intact while the missense variant is only present in NF186 could imply a different mode of disease evolution.

Molecular Modelling

To examine if these disease-associated variants affect *NFASC* structure and paranodal junction stability, 500ns molecular dynamics simulations were performed using mutant ectodomain models based on the humanised wild-type (WT) as a template. The WT *NFASC*, as well as the two variants N130D and R359P were built from structure PDB ID 3P3Y. While the transcript in which the N130D mutant is similar to the sequence of the crystallographic structure used in this study, the transcript in which the R359P mutant is different. However, for lacking of a more similar crystallographic structure and considering

that the differences are located at a considerable distance from the mutation site, the same starting structure was used to model the structure of the R359P mutant. Each protein was then placed in a cubic box and minimised.

Rearrangements on the variant sites showed that even though p.R370P gives no significant fluctuations in terms of protein conformation and movement, a weaker hydrophobic bond in residues 49 and 416 of the p.N130D protein can make its conformation less regular (Figure 12). While the WT and p.R370P settle in a stationary state, p.N130D results in a slightly compressed mutant. The weaker conformation of the protein is assessed using the RMSD, RMSF, total energy, radius of gyration and intramolecular hydrogen bonds analysis performed in comparison to the wild type protein. These scores have been established as important parameters for protein characterization. The backbone root mean squared deviation (RMSD) (Figure 13) measures the divergence of the mutant protein structure from its initial structure over the course of the simulation The radius of gyration (Rg) gives an indication on the change in size of the protein. In all cases, the most mobile portion of the chain is residue 237 (corresponding to residue 248 in the transcript containing the p.R370P variant) as seen in their root mean squared fluctuation (RMSF) which indicated how much each residue diverges form its initial position. This arginine residue is exposed to the aqueous solvent and part of a larger loop, probably flexible and crucial to protein dynamics. The RMSF further indicates that in all cases on average the overall mobility of the backbone is not particularly affected by variants.

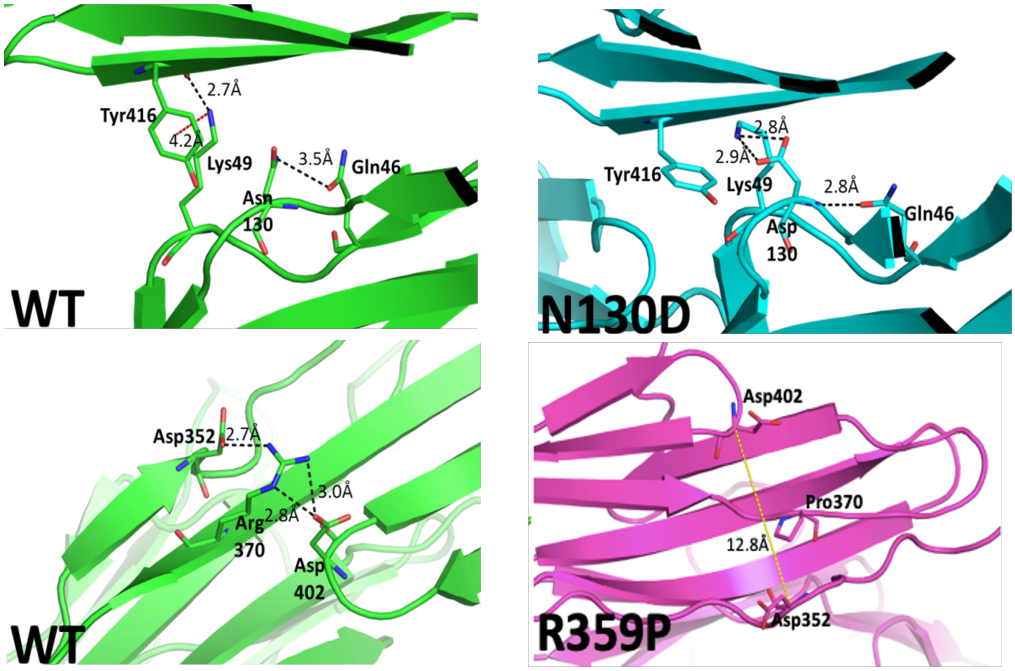


Figure 12. In silico modelling of the 3D structure of the p.N130D and p.R359P variants. In the left top panel (WT protein), a hydrogen bonding interaction with the side chain of glutamine 46 is visible. The residue of lysine 49 is distant form the point of mutation and forms a salt bridge with the carbonyl moiety of residue tyrosine 416. In addition, a weaker hydrophobic (cation-π) interaction forms between the side chains of lysine 49 and tyrosine 416 (red dashes, connecting the charged nitrogen atom with the center of the phenyl ring). The top right panel shows the different hydrogen bonding network formed in the mutant: the mutated aspartate residue forms 3 hydrogen bonding interactions, two of with involving its side chain and the side chain of residue lysine 49, and the third between its amino moiety and the side chain of residue glutamine 46. No interactions are visible between residues lysine 46 and tyrosine 416 that are further apart (about 6Å) and with an unsuitable geometry. In the left bottom panel (WT protein), a strong hydrogen bonding network connects residues aspartate 352, arginine 370 and aspartate 402 positioned on 3 different β-strands. The absence of the arginine residue in the mutant (right bottom panel) impedes the formation of the hydrogen bonding network, distancing the 3 β-strands (Efthymiou *et al.*, 2019, data generated by Sara Fortuna/Rita de Zorzi).

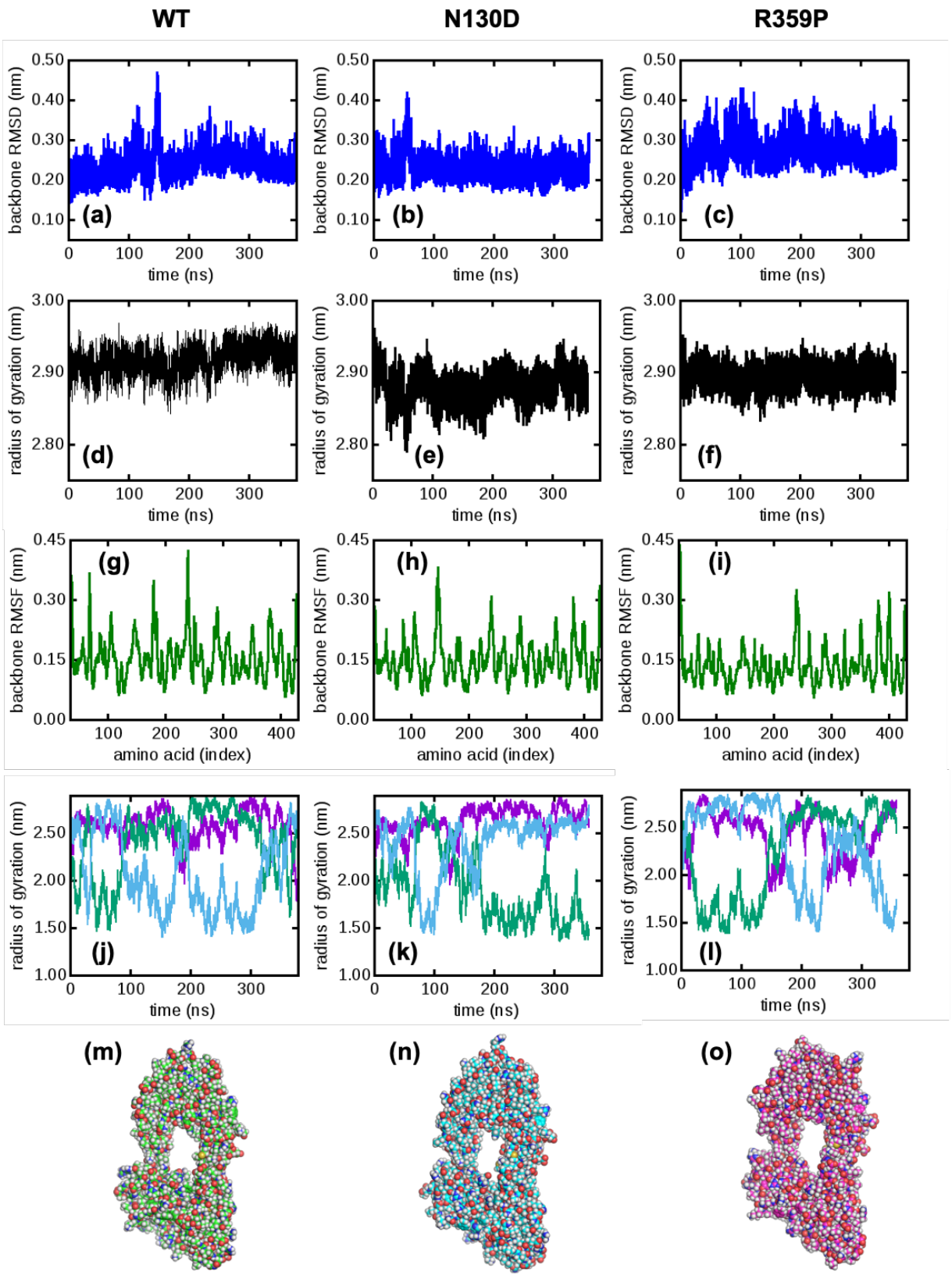


Figure 13. Root mean squared fluctuation analysis. (a-c) Backbone Root Mean Square Deviation (RMSD, blue), (d-f) backbone radius of gyration (black), and (g-i) backbone RMSF (green) for WT Nfasc186 (a,d,g) and its mutants (b,e,h) N130D, and (c,f,i) R359P. Principal axes of the neurofascin and mass weighted radius of gyration x- y- z-

components (purple, green, cyan), for (j) WT Nfasc186 and its mutants (k) N130D and (l) R359P. (m-o) Rearrangement of the Nfasc186 after 200 ns of molecular dynamics simulation for the WT protein and its mutants (Efthymiou *et al.*, 2019, data generated by Sara Fortuna/Rita de Zorzi).

Surface expression and protein levels of NFASC variants

In order to determine the impact of genetic variants on *NFASC* function, mutations were introduced in plasmids encoding Myc-tagged human Nfasc155 or Nfasc186. Protein tags such as myc and GFP are peptide sequences that are attached to proteins to facilitate easy detection and purification of expressed proteins. In the case of the p.P939Ter variant, the Myc-tag was inserted before the premature stop codon in the aberrant C-terminal extension. Human embryonic kidney (HEK) 293 cells were then transfected with these constructs and the neurofascin cell surface expression was monitored by immunocytochemistry (Figures 14 and 15). To detect surface expression of *NFASC* isoforms, living transfected cells were incubated with antibodies to Nfasc155 or Nfasc186 for 20 minutes, then cells were fixed, permeabilized and immunostained for Myc. WT Nfasc155 (Figure 14) and Nfasc186 (Figures 15 and 16) are readily expressed at the cell surface and a complete co-localization is observed between the cell surface proteins and the Myc-labelled proteins after permeabilization. All of the *NFASC* missense variants did not affect the surface expression of Nfasc155 or Nfasc186, and a clear signal of co-localization could be detected between surface-labelled and Myc-labelled proteins (Figures 14B-E and 15). Only the p.P939Ter variant strikingly affected the expression of Nfasc155. This mutant protein was not detectable at cell surface, or after Myc labelling. In order to confirm this observation, cells were co-transfected with GFP then fixed and stained for Myc. When cells were co-transfected with WT Nfasc155, all GFP-labelled cells also co-expressed Myc-tagged Nfasc155 at the cell surface (Figure 14F). By contrast, no cell

surface immuno-staining was detected in cells co-transfected with p.P939Ter (Figure 14G). This suggested that the variant abolished Nfasc155 expression.

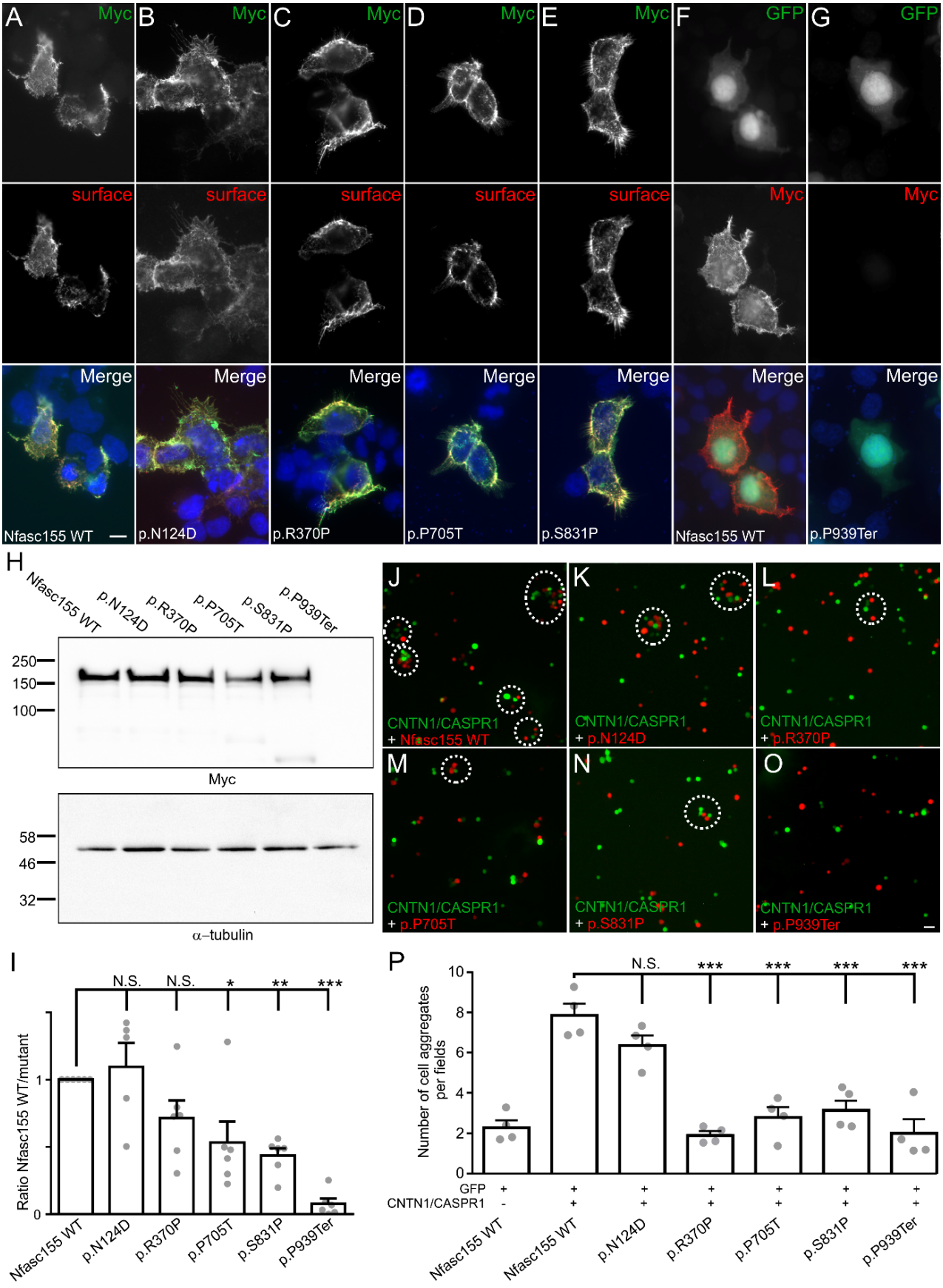


Figure 14. Membrane targeting of Nfasc155 variants. The homozygous Nfasc155

variants associated with a severe pathology diminish Nfasc155 protein level. The variants p.R370P, p.P705T, p.S831P and p.P939Ter inhibit the association of CASPR1/CNTN1 with Nfasc155. A-E. HEK293 cells were transfected with Myc-tagged Nfasc155 variants and surface Nfasc155 was monitored by incubating the live cells with anti-Nfasc155 IgG (red) prior to fixation and permeabilization. Nfasc155 was then revealed using an anti-Myc (green). WT Nfasc155, but also the variants p.N124D, p.R370P, p.P705T and p.S831P were readily targeted at the cell surface, and did not show signs of intracellular retention. F-G. HEK293 cells were co-transfected GFP (green) and either WT Nfasc155 or p.P939Ter variant, then cells were immunostained for Myc (red). By contrast to WT Nfasc155, p.P939Ter was not detectable, indicating that this mutation strongly affects Nfasc155 expression. Nuclei are stained DAPI (blue). Scale bar: 10 µm. H. Western blot analysis of HEK293 cells transfected with Nfasc155 variants and revealed with anti-Myc or α-tubulin antibodies as loading control. I. Protein expression levels were analyzed by normalizing the signals to the corresponding α-tubulin signal, then to WT Nfasc155 in 4 independent experiments. The expression levels of p.P705T, p.S831P, and p.P939Ter variants were significantly decreased compared to WT Nfasc155 (Mann-Whitney test). (*P<0.01; ** P<0.005; *** P<0.001; by unpaired two-tailed Student's t-tests for two samples of equal variance and by one-way ANOVA followed by Bonferroni's post-hoc tests). Bars represent mean and S.E.M. Molecular weight markers are shown on the left (in kDa). N.S. = not significant. J-O. N2A cells were transfected with GFP (green) in combination with CNTN1 and CASPR1 and were incubated for 2 hours with cells transfected with tandem tomato (red) and WT Nfasc155 (J), p.P705T (M) or pP939Ter (O). As negative control, N2A cells transfected with Nfasc155 (red) were incubated with cells transfected with GFP. P. The numbers of green and red cells aggregates (dashed line circles in J-O) per visualization fields was quantified in each condition (n = 4 experiments for each condition). Both p.R370P, p.P705T, p.S831P and p.P939Ter mutations significantly decreased aggregate formation (*** P<0.001 by unpaired two-tailed Student's t-tests and by one-way

ANOVA followed by Bonferroni's post-hoc tests). Bars represent mean and S.E.M. Scale bar: 20 μm. (Efthymiou *et al.,* 2019, data generated by Jerome Devaux).

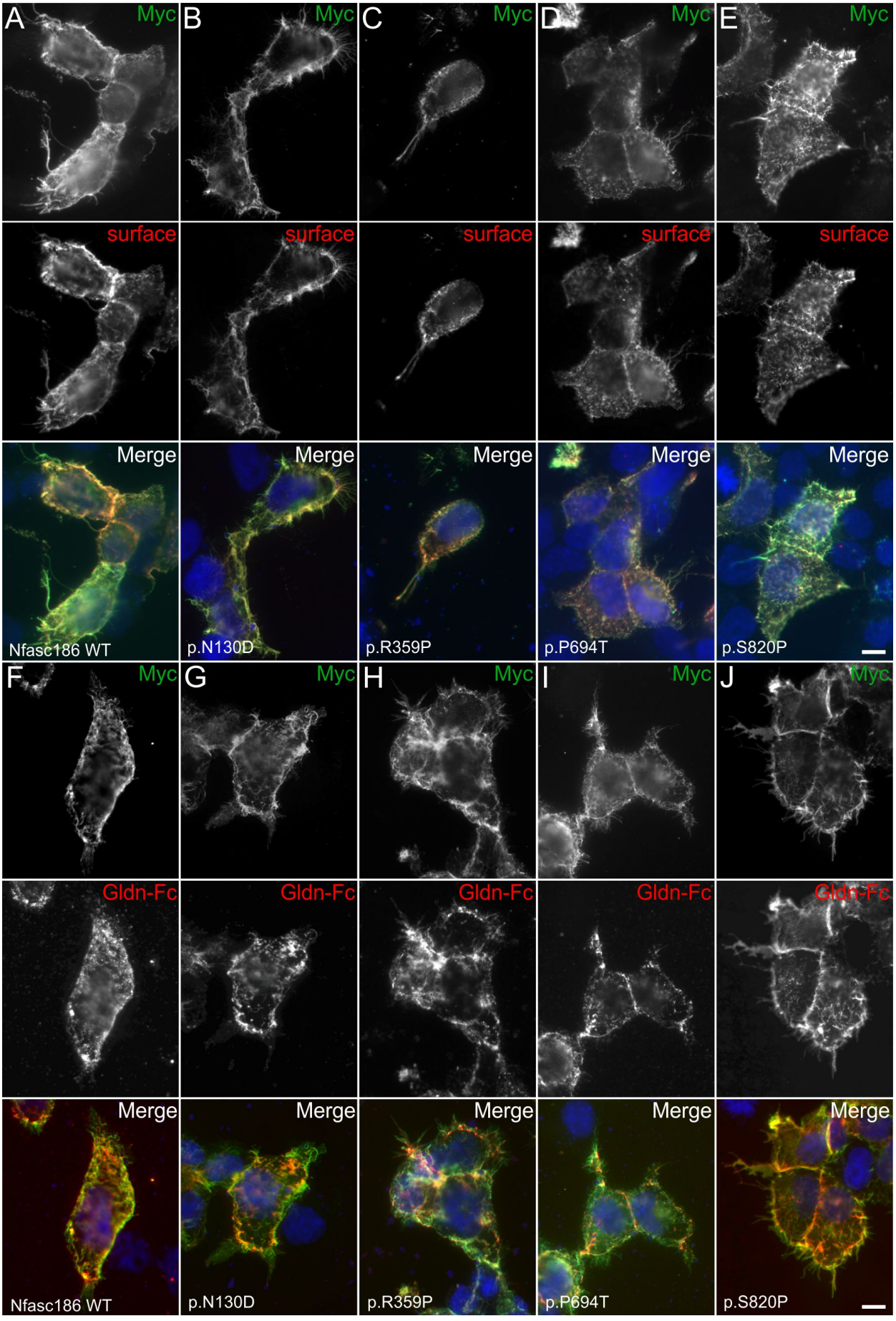


Figure 15. Surface expression of Nfasc186 variants. A-E. Myc-tagged Nfasc186 variants were expressed in HEK293 cells and their surface expression was monitored by

incubating the live cells with anti-Nfasc186 IgG (red) prior to fixation and permeabilization. Nfasc186 was then revealed using an anti-Myc (green). Nfasc186 variants appeared to be normally targeted to the cell surface. No intracellular retention was observed. F-J. Myc-tagged Nfasc186 variants (green) were expressed in HEK293 cells and the binding to gliomedin (Gldn) was monitored by incubating the live cells with Gldn-Fc proteins (1 µg; red) prior to fixation and permeabilization. Gldn bound similarly to all Nfasc186 variants. Nuclei are stained DAPI (blue). Scale bars: 10 µm (Efthymiou *et al.*, 2019, data generated by Jerome Devaux).

To further confirm this observation, we have examined the protein levels of Nfasc155 variants by Western immunoblotting (Figure 14H). The protein levels of p.P705T, p.S807P and p.P939Ter mutants were found importantly decreased compared to WT Nfasc155 (Figure 14I). Particularly, p.P939Ter was not detectable on immunoblots even at lower molecular weight. By contrast, the p.N124D and p.R370P variants had no significant effect on Nfasc155 protein level. Because, p.P939Ter is a frameshift variant inducing a premature stop codon before the transmembrane domain of Nfasc155, we suspected that this protein may be released in the culture medium. We thus harvested the cell culture medium, and immunoblotted the released fraction for Myc antibodies. No signal was detected in the released fraction even after acetone precipitation and concentration of the sample (data not shown), indicating that the p.P939Ter variant abolishes Nfasc155 expression.

Nfasc155 variants inhibit the interaction with CNTN1/CASPR1

Nfasc155 is known to interact with *CASPR1* and *CNTN1* at the paranodal regions of myelinated axons, and deletion of *CNTN1* or *CASPR1* leads to important conduction slowing in myelin-deficient mice [128]. Because the affected patients presented demyelinating features, we suspected that Nfasc155 variants may alter its interaction with

these axonal partners. To test this hypothesis, N2A cells were co-transfected with *CNTN1*, *CASPR1*, and GFP, then incubated for 2 hours with cells co-expressing Nfasc155 and tandem tomato (Figure 14). The number of cell aggregates was then quantified. As negative control, N2A cells expressing Nfasc155 were incubated with cells expressing GFP alone (Figure 14). In keeping with a previous report [135], we found that N2A cells expressing *CNTN1/CASPR1* readily interact and form aggregates with Nfasc155-expressing cells (Figure 14J-O). By contrast, only a minimal interaction was observed between Nfasc155-expressing cells and cells transfected with GFP alone (Figure 14P). Both p.P705T and p.P939Ter variants abolished the interaction between Nfasc155 and *CNTN1/CASPR1*, indicating that these variants impact Nfasc155 function (Figure 14L, O and P). To determine whether the genetic variants may also affect the surface expression of Nfasc186 or the interaction of Nfasc186 with gliomedin, we tested these on a Western blot and also tested the binding of gliomedin-Fc on Nfasc186-transfected cells. The blot showed p.N130D had a higher expression level, while the other mutants had normal expression level compared to WT Nfasc186 (Figure 15). In addition, no alterations on gliomedin binding was observed on any of the Nfasc186 variants (Figure 15).

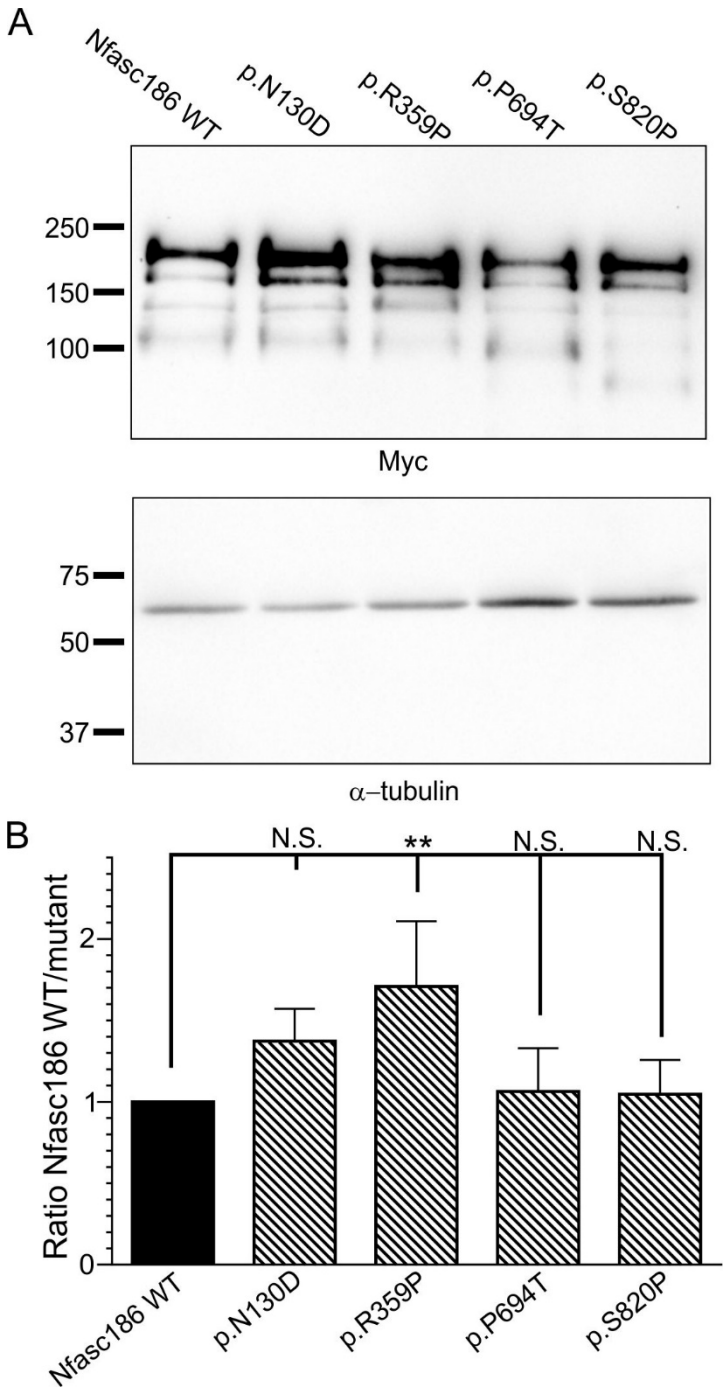


Figure 16. A. Western blot analysis of HEK293 cells transfected with Nfasc186 variants and revealed with anti-Myc or α-tubulin antibodies as loading control. Homozygous Nfasc186 variants associated with a severe pathology do not affect Nfasc186 protein level **B.** Protein expression levels were analyzed by normalizing the signals to the corresponding α-tubulin signal, then to WT Nfasc186 in 4 independent experiments. While p.R359P had a higher expression level, the expression levels of p.P694T, p.S820P and p.N130D were normal compared to WT Nfasc186 (Mann-Whitney test). (** $P<0.005$; by unpaired two-tailed Student's *t*-tests for two samples of equal variance and by one-way

ANOVA followed by Bonferroni's post-hoc tests). Bars represent mean and S.E.M. Molecular weight markers are shown on the left (in kDa). N.S. = not significant (Efthymiou *et al.*, 2019).

Immunofluorescence staining analysis

In protein gene product 9.5/collagen IV (PGP/COLIV) double stained sections of samples from both arm and leg of patient 1 (Figure 17), cutaneous innervation appeared preserved (B, B1 compared to A). In particular, epidermal nerve fibres as well as sudomotor and pilomotor nerve fibres appeared normally represented and regularly distributed. Conversely a severe loss of myelinated fibres was observed around the hair follicle as evidenced by the simplification of the so called "palisade-like" endings originating from the few remaining myelinated fibres (D, D1 compared to C). Myelin staining appeared very weak with some bright spots along the nerve course (D, D1). No further myelinated fibres were observed in MBP/PGP double stained sections.

However, small nerve fascicles were recognized in the deep dermis in sections double stained with Nfasc186/MBP (F, F1 compared to E), panNeurofascin/MBP (H, H1 compared to G) and *CASPR1*/MBP (J, J1 compared to I). Also in these fascicles the MBP-immunoreactivity appeared very weak or absent and again bright spots were present along the nerve course. Nfasc186 immunoreactivity was normally detected in the nodal region along few profiles of nerve fibres, as well as panNeurofascin and *CASPR1* immunoreactivity at the paranodes. In general, we observed a severe loss of myelinated fibres with a preservation of the unmyelinated ones. The abnormal immunoreactivity to MBP antibody suggests a severe myelin involvement leading to probable nerve degeneration. In contrast, the expression of Nfasc186 in the nodal regions and of panNeurofascin and *CASPR1* in the paranodal regions appeared normal, although limited

to the few remaining fibres. The observation of few paranodes stained with *CASPR1* as well as with panNeurofascin suggests that the absence of Nfasc155 is partial in patient 1.

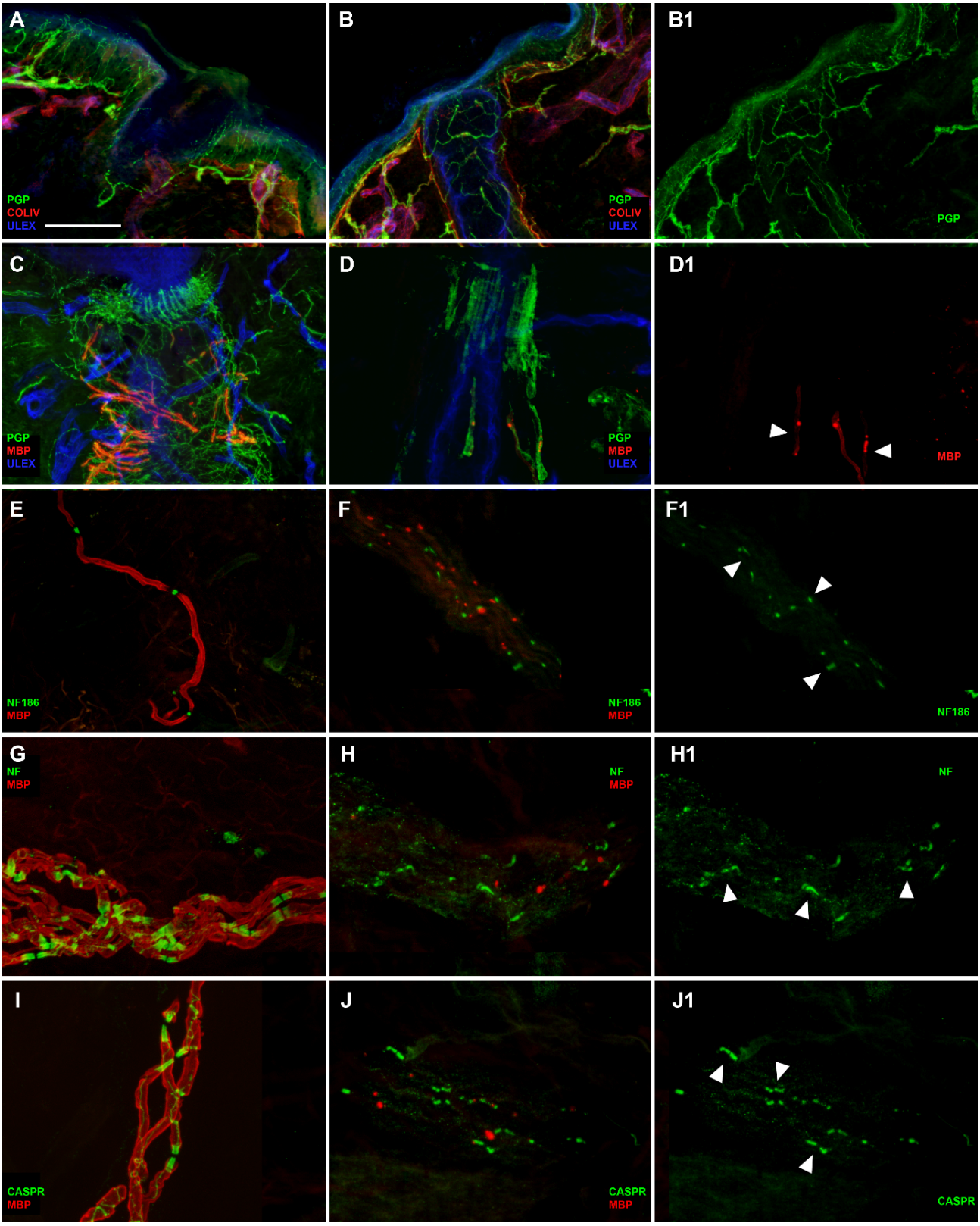


Figure 17. Severe involvement of peripheral myelinated fibers and partial disruption of Nfasc155 at the paranode in patient 1. Confocal images of cutaneous innervation from hairy skin (leg and arm) of the patient compared to a control showing: well preserved

unmyelinated fibers in the epidermis and dermis (B, B1 compared to A), severe loss of myelinated fibers around a hair follicle and the few remaining ones presenting a very weak MBP (red) immunoreactivity (D, D1 arrowheads compared to C), in a nerve bundle with few remaining myelinated fibers presenting a very faint staining with MBP antibody presence of nodes marked with Neurofascin 186 (green) (F, F1 compared to E), paranodes marked with panNeurofascin (green) (H, H1 arrowheads compared to G) and paranodes marked with CASPR (green) (J, J1 arrowheads compared to I). Bar equals 100 μm in A, B, B1, 50 μm in C, D, D1, 30 μm in F to J1. (Efthymiou *et al.*, 2019, data generated by Maria Nolano).

In PGP/ColIV double stained sections from patient 3 (Figure 18), cutaneous innervation appeared preserved in the epidermis (Figure 18A, A1) and around sweat glands and other autonomic annexes (Figure 18C) in all the 3 examined sites. Observing MBP/PGP double stained sections from distal leg, we observed a severe loss of myelinated fibres surrounding the hair follicle with few segments showing severe myelin abnormalities (Figure 18B, B1). In MBP/PGP double stained skin sections from the proximal site (thigh), myelinated fibres appeared more preserved with several myelinated segments visible around a hair follicle (Figure 18C, C1). Similarly, we could sample a definitely larger amount of myelinated fibres from the fingertip (Figure 18D, D1) compared to the other 2 sites, although there was an evident loss of mechanoreceptors and myelinated fibres compared to a control case (Figure 18E, E1). At higher magnification, we observed abnormalities of myelin sheet and enlargement of the nodal gap (Figure 19A). In Nfasc186/MBP images, we observed a normal expression of NF staining in the node (Figure 18B). The nodal structure appeared identical by visualizing the node with anti-panNeurofascin or anti-Nfasc186 (Figure 19C).

CASPR1 immunoreactivity was completely absent in the paranodal regions (Figure 19D). Additionally, an apparently normal distribution of voltage-gated sodium (Nav) channels appeared in MBP/Nav double staining images in several nodes that appeared however larger and less compacted than in the control skin (Figure 19F). Moreover, a granular distribution of Nav was evident along the axonal profiles of large fibres devoid of myelin (arrows in Figure 19E) that suggested a Nav remodeling after demyelinating phenomena. Those data indicate that the frameshift variant in patient 3, which induces a complete lack of Nfasc155 expression, does not provoke a complete disruption of the node. There was however a severe alteration of the paranode as showed by the complete absence of the *CASPR1*, the physiologic ligand of Nfasc155, and a severe effect on myelination and/ or maintenance of myelin and survival of large fibres, which appears to be length dependent. Interestingly, small fibres, both somatic and autonomic appeared less affected by the variant in both patients.

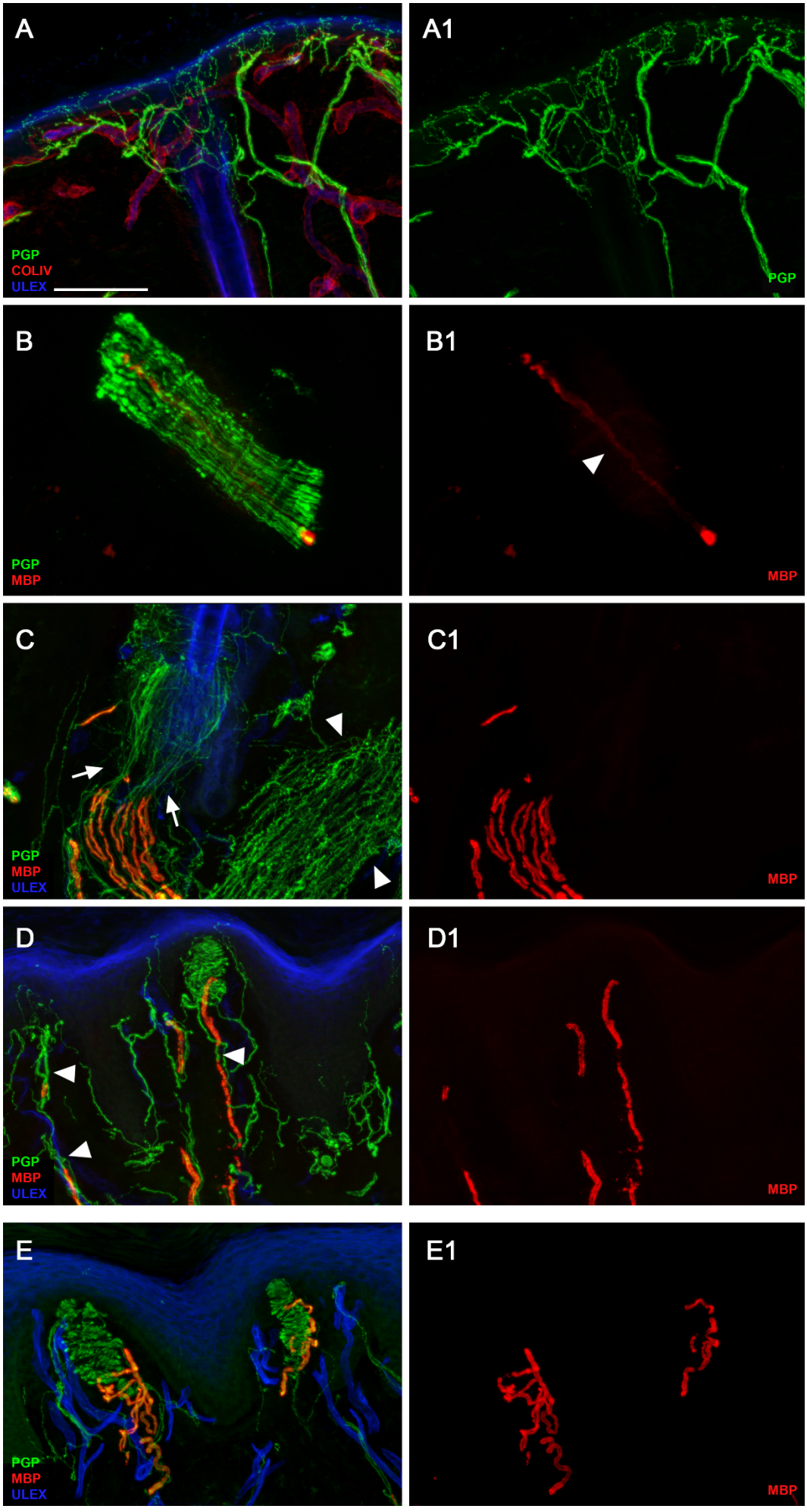

A
PGP
COLIV
ULEX
A1
PGP
B
PGP
MBP
B1
MBP
C
PGP
MBP
ULEX
C1
MBP
D
PGP
MBP
ULEX
D1
MBP
E
PGP
MBP
ULEX
E1
MBP

Figure 18. Severe involvement of peripheral myelinated axons in patient 3 with complete lack of Nfasc155 expression at paranode. Confocal images of cutaneous innervation from hairy and glabrous skin (thigh and fingertip) of the patient compared to a control showing: quite preserved unmyelinated fibers in the epidermis (A, A1) and around autonomic annexes (see arrector pili muscle, arrowheads in C); severe loss of myelinated fibers with evidence in a nerve fascicle of few segments with severe myelin abnormalities in the distal site (leg) (B, B1); more preserved myelinated fibers with evidence around a hair follicle of several myelinated fibers showing loss of myelin in their distal segments in the more proximal site (thigh) (C, arrowheads in C1); loss of mechanoreceptors and myelinated fibers (compared to the control in E, E1) and morphological abnormalities of myelinated fibers with tracts of demyelination in the fingertip (D, arrowheads in D1). (compared to the control in E, E1). Bar equals 100 μm in A, A1, C, C1, D, D1, E, E1 30 μm in B, B1. (Efthymiou *et al.*, 2019, data generated by Maria Nolano).

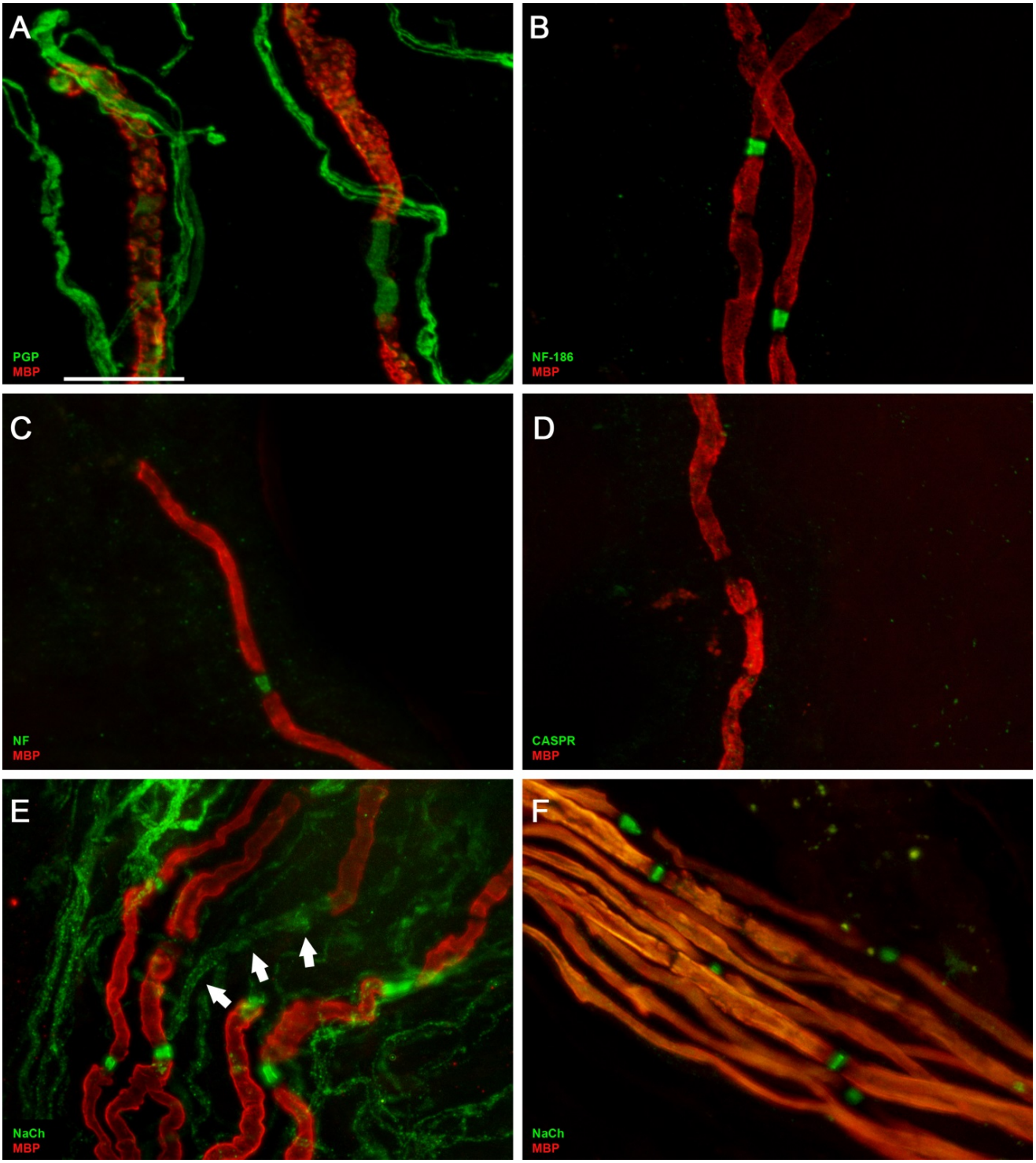


Figure 19. Paranodal alterations in myelinated axons from patient 3. Confocal images of cutaneous innervation from hairy skin of patient 3 showing abnormalities of myelin sheet and enlargement of the nodal gap (A), normal expression of Nfasc186 staining in the node as evidenced by staining with anti-Nfasc186 (NF186; B) and anti panNeurofascin (NF; C), but absent Nfasc155 and CASPR1 immunoreactivity at the paranodal regions (B and D). The distribution of Nav channels (NaCh) was normal in several nodes that appeared however larger, less compacted than in the control skin (F)

with a granular distribution of Nav evident along the axonal profiles of large fibres devoid of myelin (E) that suggests a Nav remodeling after demyelinating phenomena (Efthymiou *et al.*, 2019, data generated by Maria Nolano).

NFASC cell-specific expression and co-expression analysis

We first explored the expression of *NFASC* in the central and peripheral nervous systems using publicly available cell-specific transcriptomic data. [136; 137] In keeping with the known function of *NFASC*, RNA sequencing of purified cell types derived from human cortex demonstrated that *NFASC* was most highly expressed in oligodendrocytes though there was also evidence of expression in neurons and to a lesser extent astrocytes and endothelial cells (Figure 20).

We also found that *NFASC* expression was highest in newly formed oligodendrocytes rather than myelin-forming or mature oligodendrocytes. Next, we investigated the function of *NFASC* using gene co-expression analysis to construct networks for all 13 human brain regions sampled by the Genotype Tissue Expression Consortium (Supplementary Table 1). In the anterior cingulate cortex, *NFASC* was located with high confidence (module membership = 0.74) within a co-expression module which is enriched for genes involved in synaptic transmission (GO:0099536, chemical synaptic transmission FDR-corrected p-value = 6.68×10^{-10}) and expressed specifically in neurons (neuron module in cortex, FDR-corrected p-value = 6.0×10^{-72}). Furthermore, this module was enriched for genes known to be associated with monogenic forms of intellectual disability (FDR-corrected p-value 2.12×10^{-3}, Figure 21).

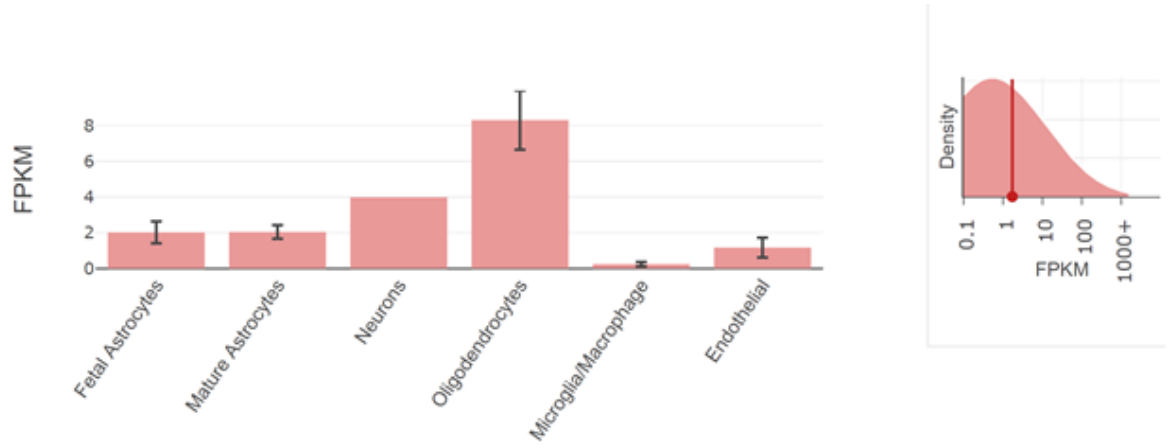


Figure 20. Cell-type specific expression of *NFASC* in human cortex generated using immunopanning [136] (http://www.brainrnaseq.org/)(Efthymiou *et al.*, 2019, data generated by Mina Ryten).

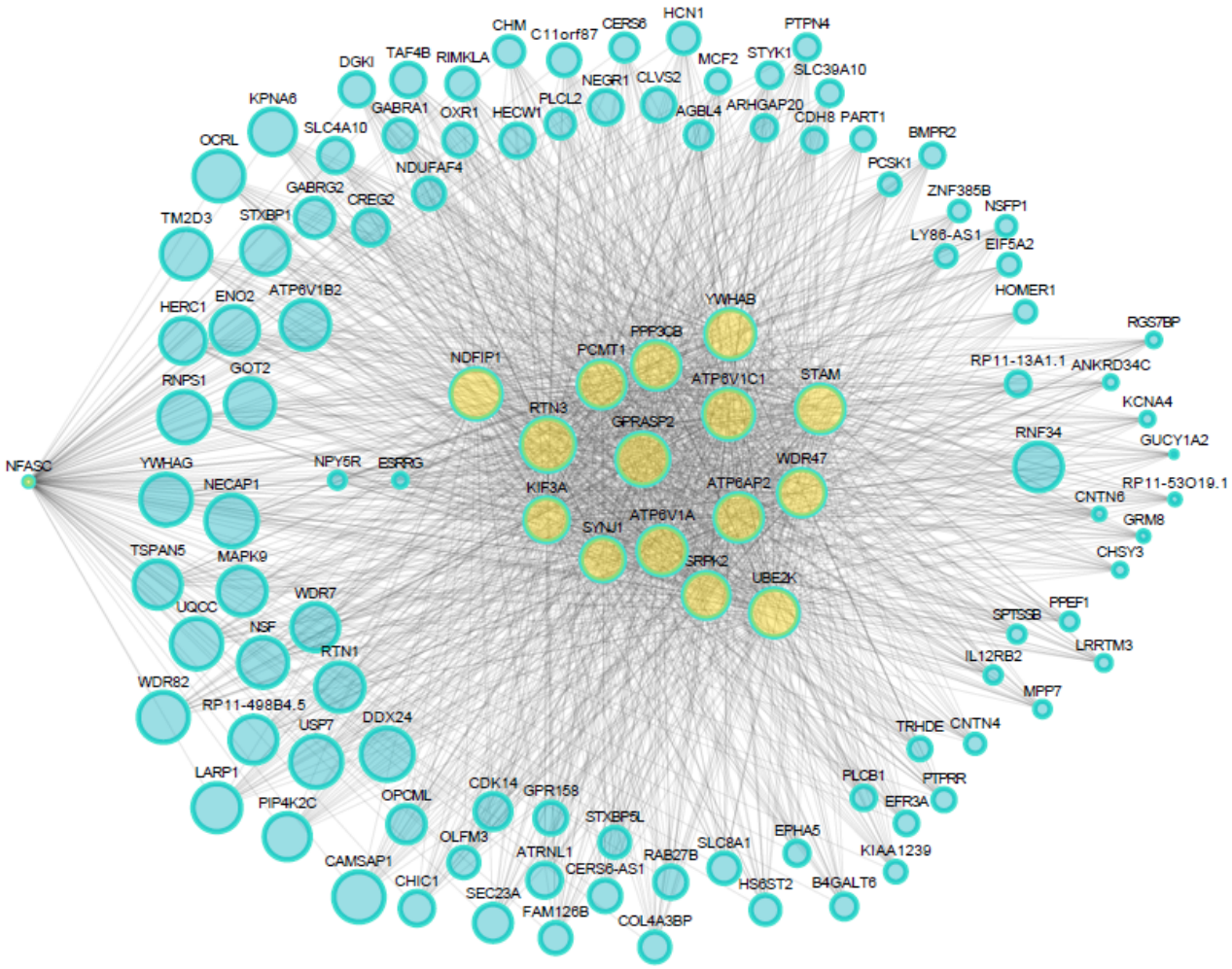


Figure 21. Cell-specific, co-expression network and brain region expression analysis plots. This module (turquoise) in the anterior cingulate cortex is significantly enriched for mental retardation/intellectual disability (ID) genes. 15 of these genes (in

yellow) together with *NFASC* (in yellow too) are used as seeds for all other genes within the module connected according to the highest strength to the seeds (Efthymiou *et al.*, 2019, data generated by Mina Ryten).

3.2.2 Discussion

Homozygous variants in *NFASC* lead to a neurodevelopmental disorder that includes in some families a chronic demyelinating neuropathy. Radiological features in some *NFASC* individuals included central white matter loss and atrophy of the corpus callosum and the brainstem. As part of their phenotype, several affected individuals also exhibited features of central demyelination and clinical and neurophysiological evidences of chronic peripheral demyelinating neuropathy associated with a severe reduction in nerve conduction velocities and prolonged motor latencies (Tables 17 and 18). Interestingly, these patients have a similar phenotype to patients with *CASPR1* mutations, however without any epilepsy which has been reported in *CASPR1*-mutated patients. It is possible that a different role is played by *CASPR1* and *NFASC* in neuron excitability at CNS level. [133; 138]

Table 18. Sensory, mixed nerve conduction and EMG studies of patient 1 carrying the p.P694T missense mutation and of patient 3 carrying the p.P939Ter frameshift mutation. Amp: amplitude, CV: conduction velocity, NO: not obtained, NA: not applicable. The parameters obtained and used for interpretation included amplitude – from baseline to peak (reflects the number of conducting fibres and is reduced in axonal loss) and conduction velocity (m/s) – calculated from the distance between stimulation and recording points, divided by latency (reflects integrity of the myelin sheath important for impulse conduction, and is reduced in demyelinating processes).

Sensory and mixed nerve conduction studies of patient 1				
	Right		Left	
	Amp (μV)	CV (m/s)	Amp (μV)	CV (m/s)
Sural (calf-ankle)	31	38	30	37
Superf. Peroneal (calf-ankle)	13	35	NO	NO
Median (upper limb-index finger)	42	48	NA	NA
Ulnar (upper limb-little finger)	29	46	NA	NA
Motor nerve conduction studies				
	Amp (mV)	CV (m/s)	Amp (mV)	CV (m/s)
Median (below elbow-wrist)	4.7	36	10.2	41
Ulnar (below elbow-wrist)	6.3	53	8.5	52
Common peroneal (fibular neck -ankle)	1.1	36	2.4	40
Tibial (ankle)	9.4	36	9.9	37

Mixed nerve conduction studies and EMG of patient 3				
Motor nerve conduction studies				
	Amp (mV)	CV (m/s)	Amp (mV)	CV (m/s)
Median (below elbow-wrist)	5.8	18.2	NA	NA
Common peroneal (fibular neck - ankle)	1.1	17.4	NA	NA
Concentric needle EMG sampling in 2008 of the deltoid, anterior tibialis and rectus femoris muscles showed acute denervation changes with increased insertional activity, fibrillations and positive sharp waves and chronic denervation with large, unstable motor unit potentials with decreased recruitment. Further sampling of the anterior tibialis muscle showed chronic denervation changes with large motor units and some polyphasia.				

Immunoglobulin G subclass 4 (IgG4) antibodies to Nfasc155 and Nfasc186 have been identified in 3-18% of patients with chronic inflammatory demyelinating polyradiculoneuropathy. [139-141] Nfasc155 IgG4 antibodies have been associated with an early onset of the CIDP-like disease, and clinical features such as congenital distal weakness, the presence of infantile sensory ataxia and tremor [139; 140; 142] were also seen in some of our patients. A cerebellar origin of the tremor has been hypothesized in some cases. Also, and most importantly to the present study, up to 8% of CIDP patients with IgG4 Nfasc155 showed evidence of CNS demyelination [143] and anti-Nfasc155 were identified at high frequency in Japanese patients with combined central and peripheral demyelination [144] , although this was not replicated in the general population. [145]

The majority of *NFASC* mutated individuals identified in this study were found with homozygous missense variants affecting highly conserved residues within the Ig-like domains or in the Fibronectin type III domains of the neurofascin protein (Figures 22 and 23). One family (Family 2) was identified with a homozygous frameshift deletion predicted to result in a premature truncation of the protein. Among our patients, affected individuals from Family 2 carrying the p.P939Ter frameshift deletion exhibited the more severe phenotype with neurodevelopmental impairment and profound muscle weakness associated with severe reduction of conduction velocities on neurophysiological investigations. Given the history of consanguineous marriages in these families, other recessive variants may also have contributed to the clinical phenotype, as in proband 8 from Family 5 who died prematurely presumably because of aspiration due to respiratory difficulties.

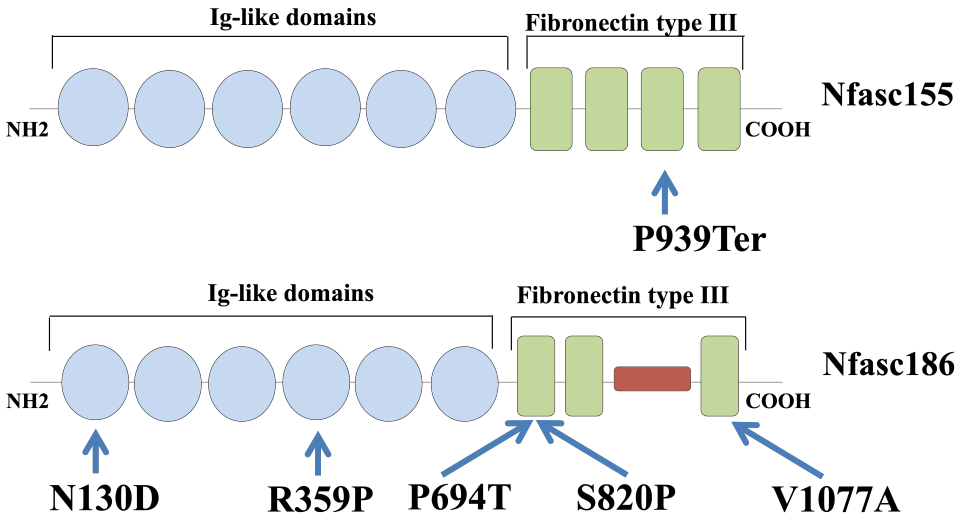


Figure 22. A schematic representation of Nfasc155 and Nfasc186. It shows the position of all *NFASC* variants (Efthymiou *et al.*, 2019).

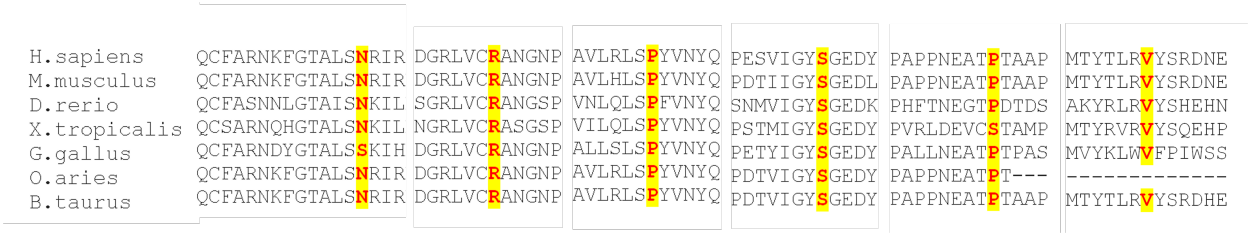


Figure 23. Interspecies alignment performed with Clustal Omega. It shows the complete conservation down to invertebrates of the amino acid residues affected by the substitutions (Efthymiou *et al.*, 2019).

The results from our functional analyses indicate abnormal Nfasc155 interaction with CNTN1 and CASPR1 as an important disease mechanism in the *NFASC*-related genetic disease. In addition, loss of surface expression may also be implicated in the phenotype observed in the affected individuals. Immunofluorescence studies on patients 1 and 3 detected a severe loss of myelinated fibres with a preservation of the unmyelinated ones, and while the expression of Nfasc186 in the nodal regions appeared normal in both patients, the absence of Nfasc155 in the paranodal region was only partial in patient 1. Western blot showed that the protein levels of several Nfasc155 mutants including p.P705T and p.P939Ter were importantly decreased compared to WT Nfasc155. These findings supported the morphological findings and indicated that there may be an

incomplete derangement of the paranodal region in some patients due to a reduction of glial Nfasc155. This could explain the relatively milder phenotype of patients 6 and 7 as well as that of the recently described patient with a similar neurodevelopmental syndrome.[2]

Of interest, in cell-specific *in silico* studies NFASC is mostly expressed in newly formed oligodendrocytes, and found to be enriched for genes associated with monogenic forms of intellectual disability. Given recent evidence, that the generation of new myelin segments is an important form of plasticity that can be used to modify the properties of circuits in the CNS[146; 147], these observations may be significant and provide some explanation for the neurodevelopmental features of individuals carrying biallelic *NFASC* variants. We could further speculate that in at least some brain regions *NFASC* may play an important role in synapse transmission and that this potentially might explain how biallelic variants in the gene result in a phenotype which includes intellectual disability.

Recently, a homozygous splice-site variant in *NFASC* has been identified in an individual with neurodevelopmental impairment, hypotonia and areflexia.[148] Another family was found within a large genetic study carrying a homozygous *NFASC* variant associated spinal muscular atrophy [2] and an homozygous missense mutation leads to significant loss of *NFASC* in iPSC-derived neurons from affected subjects.[4]

Despite these reports, the role of pathogenic *NFASC* variants in disease development remains to be elucidated and warrants further investigations. Our results strongly establish the association of biallelic *NFASC* variants in the complex central and peripheral nervous system involvement presented by the 8 affected individuals described in this study. The notion of a genetic *NFASC* disease further delineates an emerging spectrum of human neurological disorders caused by variants in cell-adhesion and ankyrin-binding genes

involved in neurite extension, axonal guidance, synaptogenesis, myelination and neuron-glial cell interactions.

Interestingly, homozygous *NFASC*-null animals born with a normal appearance but die suddenly 7 days after birth when the postnatal transition to saltatory conduction is occurring in the CNS and PNS. Overexpression of either of the two neuronal isoforms (Nfasc186 or Nfasc140) helps them survive into adulthood but mice are ataxic, presumably because they do not have intact paranodal axoglial junctions in their myelinated nerves.[128]

3.2.3 Further work. CRISPR-aided genome editing in mouse.

I have successfully applied and in late 2019 acquired a seeding grant of 20,000 Euros from the "SOLVE-RD – Solving the Unsolved Rare Diseases" project, funded from the European Union's Horizon 2020 research, to generate and characterise a genetically altered *NFASC* mouse mutant. The generation of point mutations within *NFASC* will be reached by using CRISPR/cas9 technology. This will be carried out in collaboration with the MRC Harwell team (Dr. Sara Wells, Dr. Lydia Teboul, Dr. Michelle Stewart).

We have designed the following outlined project:

- Aim 1 (Month 1): Design of molecular biology reagents
- Aim 2 (Month 2): Delivery of reagents into pre-implantational mouse embryos
- Aim 3 (Month 4): Quality control of founder mice and mating
- Aim 4 (Month 6): Confirmation of germ line transmission, QC and further breeding
- Aim 5 (Month 9): Breeding for phenotyping
- Aim 6 (Month 12-18): Comprehensive phenotyping
- Aim 7 (Month 12-18): Mouse Tissue will be sent to UCL lab for further analysis

At the end of this project, MRC Harwell and UCL will jointly draft a publication (if applicable) as well as report in the form of functional validation to Dr. Holm Graessner at EKUT and to the Solve-RD team.

The following experimental procedure will be carried out:

- Mutation design and generation of reagents for microinjection
- Single guide RNA (sgRNA) sequence selection is carried out
- Sequences for donor templates are designed with homology arms at least 60 nt in size flanking the intended point mutation.
- In vitro RNA transcription where single-guide RNAs are synthesised
- Cas9 mRNA is either synthesised or commercially sourced
- In vitro evaluation of sgRNA efficacy
- B6N mice housed and maintained in the Mary Lyon Centre, MRC Harwell Institute used
- Pronuclear microinjections of zygotes
- Genotype screening and characterisation of genomic DNA from F0 and F1 animal ear clips
- Mating of a mutant founder F0 to a WT counterpart allow the generation of a heterogeneous F1 population of heterozygous animals, all containing one WT allele and one allele to characterise (Figure 24).
- PCR amplification and Sanger sequencing of all F1 animals generated for the point mutation, and only those with it should be performed to establish the exact genetic makeup of a new mutant line (homozygous population representing the inheritance of the gene of interest).
- Phenotyping using tools such as SHIRPA, grip strength, startle response (Figure 25).

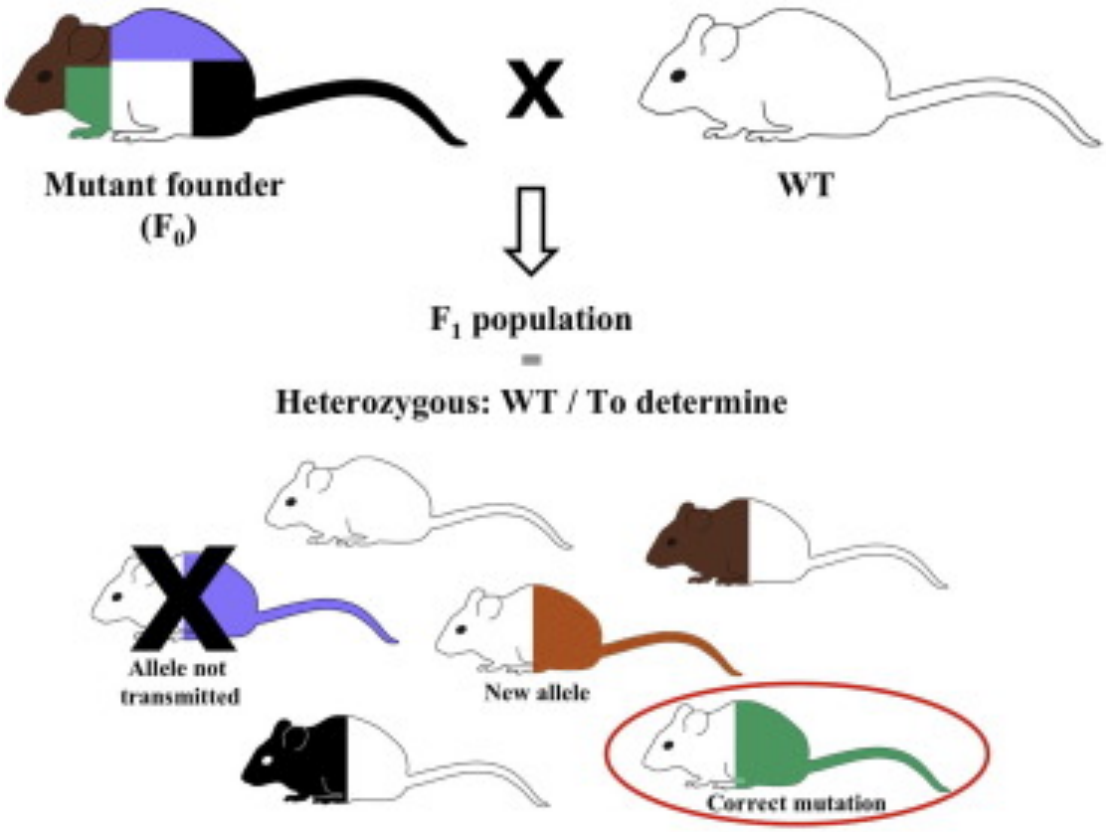


Figure 24. Mosaicism and germ line transmission complexity. Mating of a mutant founder F0 to a WT counterpart allow the generation of a heterogeneous F1 population of heterozygous animals, all containing one WT allele and one allele to characterise (Miannée *et al.*, 2017).

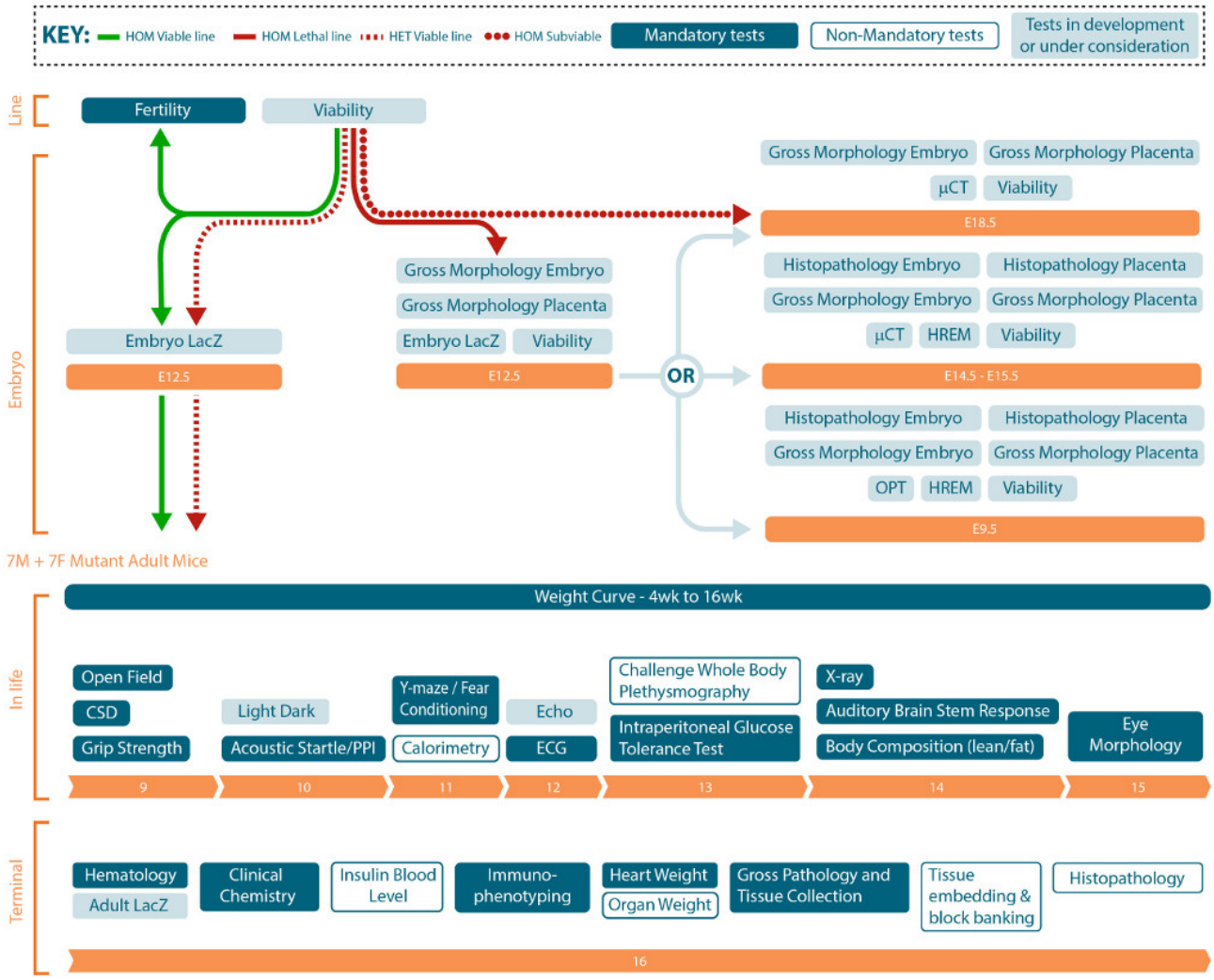


Figure 25. The IMPC (International Mouse Phenotyping Consortium) core pipeline outlining the phenotype pipeline that has been developed by all IMPC phenotyping working groups (from https://www.mousephenotype.org/).

3.3 Human patient *SFPQ* homozygous mutation is found deleterious for brain and motor development in a zebrafish model

3.3.1 Introduction

The development of the human brain relies on complex mechanisms involving both the generation of appropriate cell types and their organization in the correct anatomical structures. Post-transcriptional regulation of gene expression plays a fundamental role in the temporal and spatial modulation of early development and has been recently found to be involved in many disease mechanisms.

Cerebellar dysfunction is a recurrent feature of several neurodevelopmental disorders, including intellectual disability (ID), autism spectrum disorders (ASDs), developmental epileptic encephalopathies (DEEs) and speech disturbances, and abnormal cerebellum development (during early embryonic stages) can significantly contribute to the anomalies of movement, cognition, and affective regulation frequently observed within the phenotypic spectrum of these disorders.[149] Mono and biallelic mutations in genes implicated in RNA processing and metabolism can affect critical post-transcriptional events from the earliest stages of life.[150] Understanding the impact of mutations in such genes on neurodevelopment and disease phenotypes will allow to elucidate initiating molecular events as well as guide the development of new therapies targeting primary mechanisms before the disease progresses too far.

SFPQ encodes a Splicing Factor Proline/Glutamine-Rich (Sfpq) multifunctional protein, ubiquitously expressed across the body from embryo to adult in all vertebrates and known to be involved in transcriptional regulation, pre-mRNA splicing and RNA transport. [151-154] Human *SFPQ* lies in a region on chromosome 1p34-p36 already implicated in speech anomalies and language impairment and has been found to be mis-regulated in brains of

patients with a variety of neurodevelopmental disorders such as autism and dyslexia. [155-158] The SFPQ protein is conserved across species and plays a key role in neuronal development and network organization.[159] The protein structure contains tandem RNA recognition motif domains, a NOPS domain, a coiled-coil region and an N-terminal proline/glutamine-rich low-complexity region.[160] More recently, loss of SFPQ function has been implicated as a risk factor for human neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), fronto-temporal dementia (FTD) and Alzheimer's Disease (AD) in human tissue, mouse, iPSC-derived neurons, and zebrafish models. [83; 161-163]

Antibody staining in zebrafish showed that the protein is localised in nuclei in all tissues throughout development and in neurites of a subset of neuronal populations, including motor neurons.[159] This unique cytoplasmic pool was found to drive motor neuron maturation in absence of nuclear function.

Recent RNA sequencing studies of patient iPSC-derived neuron cultures from both familial and sporadic ALS patients showed abnormal premature intron retention in the *SFPQ* transcript during early neural differentiation across the mutant cultures, leading to loss of function of SFPQ. These results revealed SFPQ loss of function as an early hallmark of ALS.[161]

In this study, we identified a homozygous p.Ser660Asn variant in *SFPQ* through whole exome sequencing (WES) of an Italian family affected with a complex neurological phenotype characterized by intellectual disability, peripheral neuropathy, vacuolar myopathy, parkinsonian features and neuroradiological anomalies resembling neurodegeneration with brain iron accumulation (NBIA).

In collaboration with Prof. Corinne Houart (King's College London) rescuing a zebrafish *SFPQ* homozygous null mutant with this human SFPQ^{S660N} variant revealed robust defects in the developing central nervous system (CNS) of the embryos, including microcephaly, misfolding of the posterior brain neuroepithelium and abnormal branching of the motor axons innervating body muscles. The defects identified in the SFPQ^{S660N} zebrafish model provide evidence supporting a contribution of SFPQ^{S660N} to the patients' complex pathology and suggest that *SFPQ* mutations may contribute to different broad clinical and neuroradiological degenerative features in humans.[164]

3.3.2 Results

Clinical findings

We report the finding of a link between a severe genetic neurodevelopmental disorder and a homozygous missense mutation in *SFPQ*. We investigated the exome of a female patient presented with microcephaly, neuropathy, head tremor and severe cerebellar atrophy with nystagmus and dysarthria. The clinical data was provided by Dr. Carmela Scuderi. The patient was a 47-year-old woman, born at term by normal delivery from healthy non-consanguineous Sicilian parents (Figure 26). The patient's family history was positive for intellectual disability in her first-degree cousin. Her birth weight, length and Apgar scores have not been reported. Since the first months of her life, the patient presented muscle hypertonia and a severe delay in psychomotor development with poor spontaneous movements, lack of head control, frontal plagiocephaly, weak crying, and absent suction. An episode of seizures triggered by fever was reported at the age of 1 year. One year later, she presented with an episode of fever-free generalized tonic-clonic seizures. From early adulthood, she developed recurrent partial epileptic episodes and treatment with valproic acid was started, with partial control of seizures. These episodes were frequently accompanied by staring and brief loss of consciousness. The patient was first admitted to the IRCCS Oasi Maria SS. Troina Medical center (Troina, Italy) at the age

of 38 years old. At this time, physical examination revealed microcephaly, congenital absence of teeth, flaccid abdomen (with accumulation of fat or skin), and skin trophic alterations. Neurological examination revealed spastic tetraparesis and ataxia with ocular motor apraxia, also featuring marked deficits of horizontal eye movements and nystagmus. At this time, she also showed signs of bradykinesia, extrapyramidal rigidity (with abnormally increased resistance to movement) and rest tremor, mainly involving the head. Babinski reflex was absent and also deep tendon reflexes in the four limbs were not elicitable. Dysmetria and intention tremor were present in the upper limbs. Muscle wasting was also predominant in her lower limbs with retraction of hip adductor group and triceps surae muscles. A neuropsychological assessment as part of her neurological evaluation revealed a severe intellectual disability. Electromyography (EMG) showed motor and sensory axonal polyneuropathy.

Electroencephalogram (EEG) highlighted focal diffuse paroxysmal activity over the frontal and the parietal regions. Brain magnetic resonance imaging (MRI) performed at the age of 40 years old showed severe cerebellar atrophy, associated with brainstem atrophy, and thin dysplastic corpus callosum (Figure 27). In the supratentorial area, white matter atrophy and signal changes were also documented. A symmetric hypointensity of the globus pallidus and of the mesencephalic substantia nigra was also documented, suggesting anomalous deposit of paramagnetic substances (e.g., iron). Brain computed tomography (CT) scan did not show intracranial calcifications.

Ophthalmological and fundoscopy examinations identified bilateral optic atrophy and strabismus. Visual evoked potentials (VEPs) appeared prolonged. Electrocardiogram (ECG) and echocardiogram were normal. The abdominal ultrasound described an unspecified hepatopathy. Urine organic acids, karyotype and array comparative genome hybridization were all normal. Panel NGS analysis for 12 genes implicated in

neurodegeneration with brain iron accumulation (NBIA) did not show any pathogenic mutations or intragenic deletions/duplications. Muscle biopsy (image not available) gave a picture suggestive of vacuolar myopathy associated to areas of adipose fat replacement, neurogenic dysfunction, and denervation. Also, a reduction of respiratory chain complex I and II was documented. At the follow-up appointments, a progression of the neuromotor symptoms, featuring an increased rigidity to the passive movements and increase of the tremor, has been observed. Follow-up brain imaging studies confirmed the previous radiological findings.

Genetic findings

Whole-exome sequencing of the trio generated in total, 55,531,100 (II-1) unique reads. It revealed a heterozygous de-novo mutation (NM:004321: c.920G>A; p.Arg307Gln) affecting a conserved residue within the motor domain (MD) of the *KIF1A* gene. In addition, a homozygous missense variant in the *SFPQ* gene (NM_005066:c.G1979A; p.Ser660Asn) was also identified. The p.Ser660Asn *SFPQ* variant involves a conserved Serine residue and was present within the most significant homozygous block (chr1: 33549405-40840270) identified by homozygosity mapping analysis (which was performed using the WES data). The variant is absent from GnomAD (gnomad.broadinstitute.org) containing 125,748 exome sequences from unrelated individuals, and it is considered pathogenic by several in silico predictors (including Mutation Taster, FATHMM, Sift, LRT, Eigen, MetaLR). Both parents were found to be heterozygous carriers of the p.Ser660Asn variant in Sanger-based segregation analysis (Figure 26). The p.Arg307Gln *KIF1A* variant could partially explain the complex neurological phenotype of our patient as it is established in ClinVar (www.ncbi.nlm.nih.gov/clinvar; accession number: 418275) and in literature reports as a disease-causing mutation, implicated in phenotypes ranging between intellectual disability, spasticity, optic nerve and/or cerebellar atrophy. [165-167] Rare *SFPQ* variants have been implicated in ALS and neurodegenerative diseases [83] and, more

recently, loss of SFPQ was implicated in a peculiar skeletal myopathy in a mouse model [168].

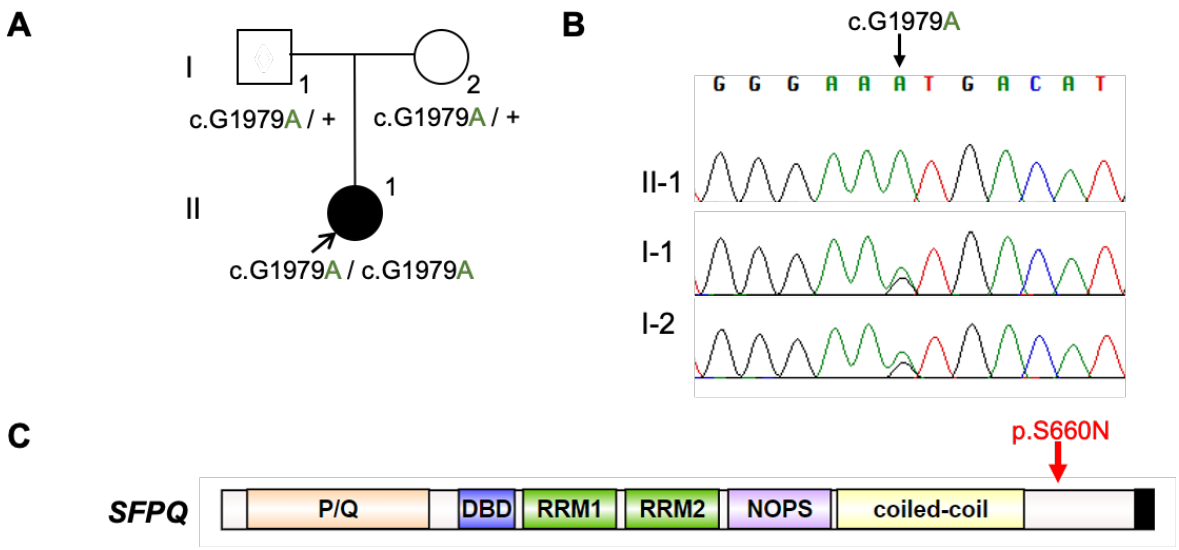


Figure 26. Pedigree and Sanger sequencing of the SFPQ variant. A. Pedigree of the Italian family showing the inheritance of the *SFPQ* variant in proband and parents. **B.** Segregation analysis of the *SFPQ* variant by traditional Sanger sequencing, showing the *SFPQ* variant identified as homozygous in the proband and heterozygous in the two parents. **C.** Protein structure of SFPQ, made up of important domains mentioned above, and showing the location of the *SFPQ* variant identified (modified from Efthymiou and Gordon, 2020).

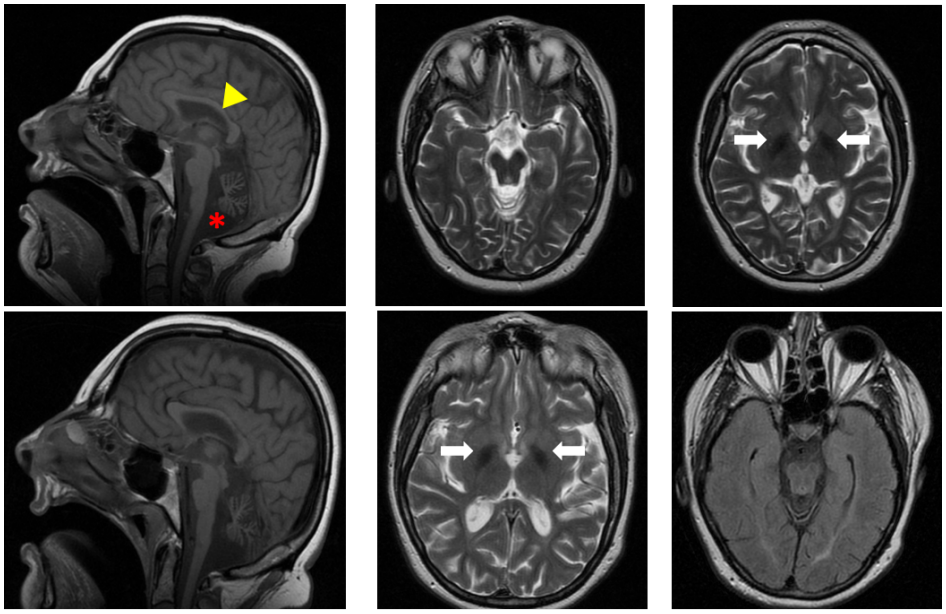


Figure 27. Radiological findings of the SFPQ patient. Top panel: MRI studies at the age of 36 years. Sagittal T1 and axial T2 studies reveal severe cerebellar and vermian atrophy, with brainstem ipotrophy (red asterisk) and thinning of the corpus callosum (yellow triangle). In the supratentorial area, periventricular white matter signal abnormalities are evident. Symmetric hypointensity of the globus pallidus is also present (white arrows). Bottom panel: At the age of 40 years abnormal white matter signal, cerebellar and vermian atrophy and brainstem hypotrophy is documented in T1 axial images. Hypointensity of globus pallidus in T2 and inversion recovery fast spin-echo flair sequences is also detected (modified from Efthymiou and Gordon, 2020).

The KIF1A heterozygous hit being insufficient to explain all aspects of the patient' disabilities, the rare homozygous *SFPQ* variant emerged as a potential candidate contributor to some of the clinical and radiological features. These include the parkinsonian features of bradykinesia and rest (head) tremor, the degree of cerebellar atrophy (more severe than what is usually observed in the disease associated monoallelic *KIF1A* variants), the history of myopathy (with widespread vacuolated muscle fibers as documented by biopsy studies), and the presence of 'NBIA-like' findings on brain imaging (Ferritin heavy chain transcripts are downregulated in zebrafish SFPQ null mutants).[83]

Zebrafish characterisation

To test the functional impact of the *SFPQ* p.Ser660Asn mutation, we used the zebrafish as a vertebrate model. We previously showed that *sfpq* null loss of function homozygous mutation leads to brain boundary defects, cerebellar abnormalities and loss of motor function in zebrafish [83] and found a link between this gene and Amyotrophic Lateral Sclerosis (ALS). We have shown that human full-length *SFPQ* is able to rescue completely the loss of gene function in zebrafish. We therefore tested whether the human SFPQ^{S660N} transcript had the ability to rescue the zebrafish homozygous null mutant and if so, whether this rescue was accompanied by neurodevelopmental abnormalities. Rescue experiments are done blind by injection of full-length wild-type hSFPQ or hSFPQ^{S660N} transcripts at 1-cell stage (SFPQ is expressed ubiquitously from fish to human) into an incross of *sfpq*$^{+/-}$ zebrafish (25% of the embryos are homozygous null), followed by pan-axonal staining at 36- or 48-hours post-fertilisation (hpf).

We first performed a set of double-blind experiments to assess whether any visible abnormalities were associated with rescue of the null mutant by the S660N allele. Homozygous mutant embryos rescued with SFPQ^{S660N} showed robust abnormalities in brain formation and motor innervation (Figure 30), so we further investigated the developing CNS in these embryos using phalloidin and anti-acetylated tubulin that labels specifically all axons in zebrafish (Figure 28). This analysis confirmed abnormal branching of the motor axons innervating body muscles (Figure 28J-L,N,O) and a spectacular misfolding of the neuroepithelium at the isthmus, bulging either into the tectal area (equivalent to superior colliculus) or into the rostral brain stem (Figure 28A-L) in addition to a reduction of telencephalic volume (Figure 28P) and of the distance between the eyes due to reduction of hypothalamic volume (Figure 28Q), without disruption of retinal organization (Figure 28M). The isthmic dysmorphology is accompanied by a lack of differentiation of the cerebellum.

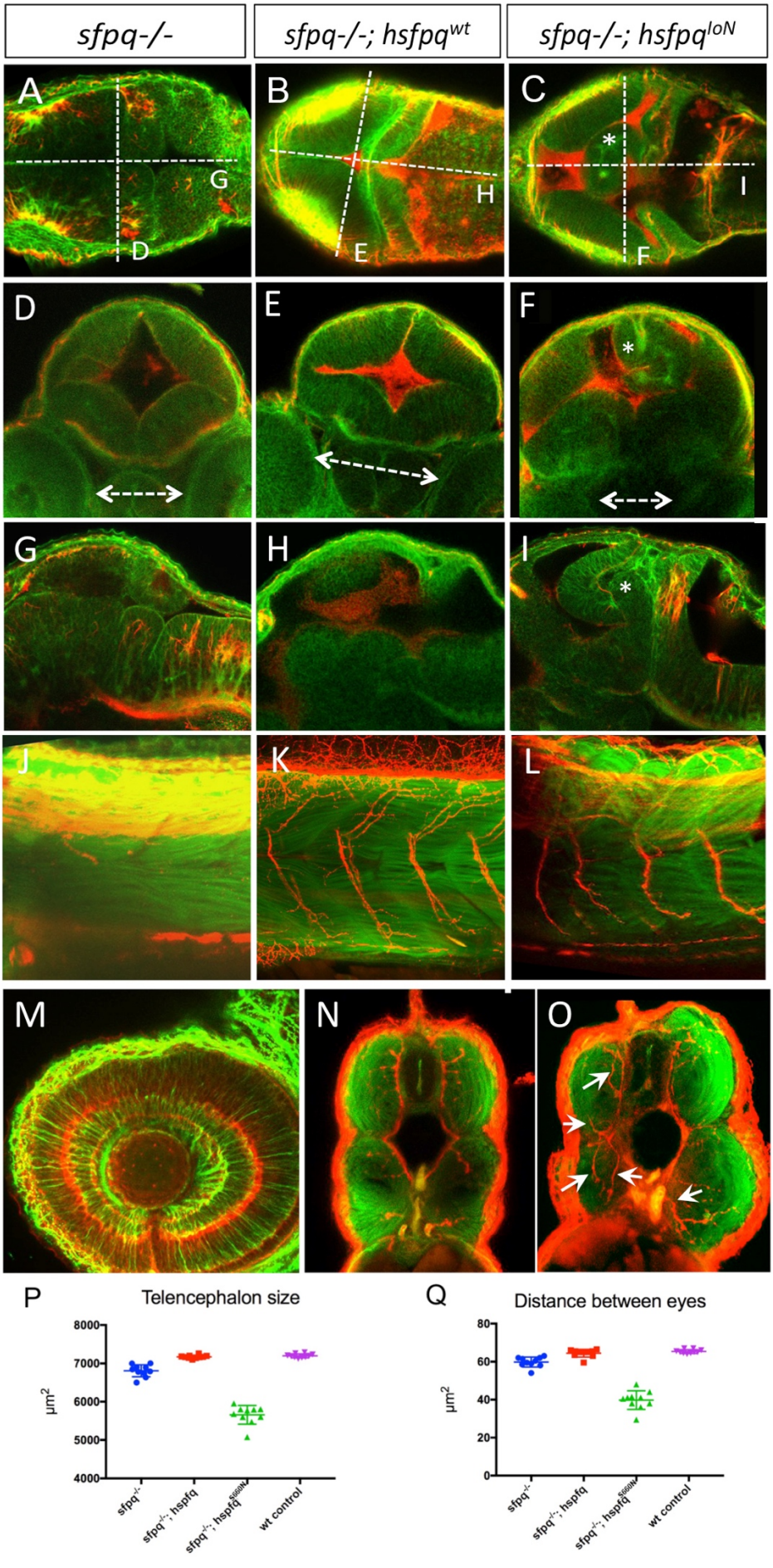

sfpq-/-
sfpq-/-; hsfpq^wt
sfpq-/-; hsfpq^loN
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Telencephalon size
µm²
8000
7000
6000
5000
4000
sfpq-/-
sfpq-/-; hspfq
sfpq-/-; hspfq S660N
wt control
Q
Distance between eyes
µm²
80
60
40
20
0
sfpq-/-
sfpq-/-; hspfq
sfpq-/-; hspfq S660N
wt control

Figure 28. Phenotypic characterisation of *sfpq-/-* zebrafish embryos rescued by ubiquitous expression of human *SFPQ*^{S660N} gene. A-I: Dorsal (A-C), transverse (D-F, double arrow showing distance between eyes), and lateral (G-I) views of the midbrain/hindbrain Isthmus at 48hpf showing dramatic misfolding (*) in embryos expressing the S660N variant (C, F, I) but not in the mutant rescued by expression of wild-type human SFPQ (B, E, H). **J-L, N, O:** Lateral views (J-L) and transverse sections (N, O) of trunk motor innervation in wild-type (K, N) or S660N variant (L, O) rescued 48hpf embryos. M: Lateral view of the eye in the S660N rescued animal, showing no defect in retinal organization. **P, Q**: Quantification of telencephalic size and optic distance measured at 48hpf in ten wild-type, (purple), *sfpq-/-* uninjected (blue), *sfpq-/-* injected with wild-type human *SFPQ* (red) or with human *SFPQ* S660N variant (green) embryos, Phalloidin staining in green and acetulated tubulin in red (Efthymiou and Gordon, 2020, data generated by Patricia Gordon).

To identify whether the misfolded tissue was of midbrain or hindbrain identity, we performed *Otx2* (midbrain marker), *Gbx2* (hindbrain marker) double *in situ* mRNA hybridization and found that the ectopic lump was of dual origin, comprising *Otx2*- and *Gbx2*-expressing cells (Figure 29).

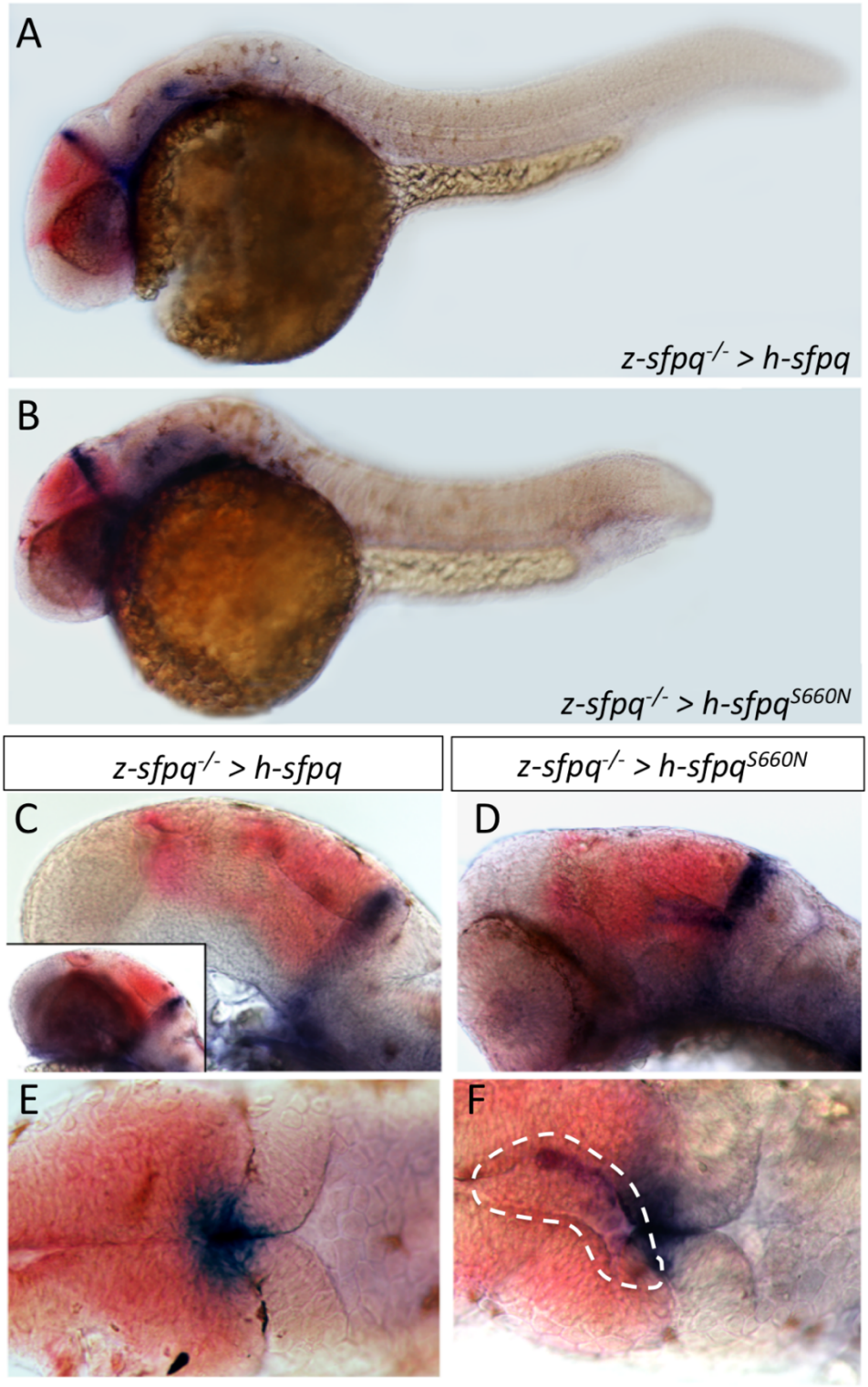


Figure 29. Tissue identity of the ectopic ventricular brain mass. A-B: Lateral view of whole sfpq-/- embryos rescued by wild-type (A) or S660N (B) human *SFPQ,* showing *Otx2* (red) and *Gbx2* (blue) expression. Note that the variant-rescued embryos have most often a shorter body. **C-F:** Lateral (C, D) and dorsal (E, F) close-up view of the midbrain/hindbrain boundary. Inset in C shows the head before eye dissection. This was not possible to achieve in the variant-rescued embryos due to greater adherence of the

retina to the rest of the brain. Dotted line in F delineate the ectopic ventricular mass (Efthymiou and Gordon, 2020, data generated by Patricia Gordon).

We also examined the spinal cord motor neurons of embryos rescued with SFPQ^{S660N} RNA and found that SFPQ^{S660N} led to dramatically shortened motor axons that exhibited ectopic branching and disorganized morphology (Figure 30). All phenotypic analyses of the zebrafish embryos were done without prior knowledge of the patient pathologies. The defects found are strikingly similar to some of the developmental problems of the patient as suggested by the MRI data.

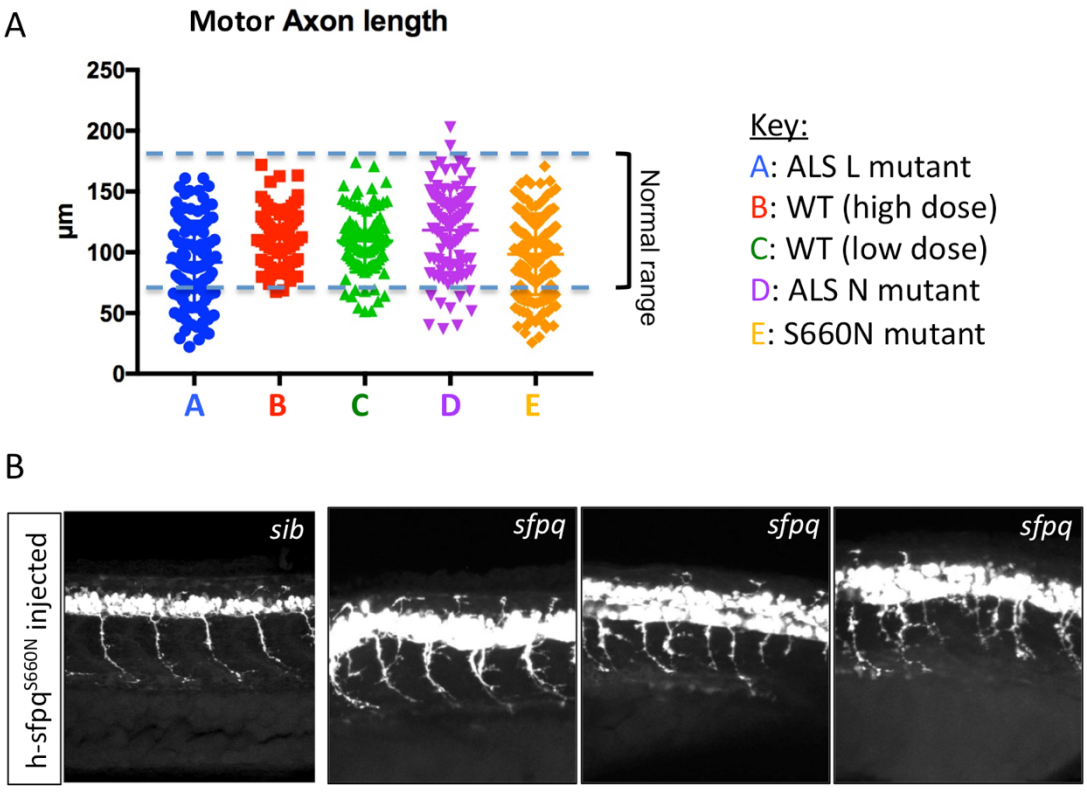


Figure 30. Double-blind phenotypic assessment of zebrafish SFPQ null mutant rescued by various human SFPQ variants

A: Length of motor axons in the progeny of *sfpq*$^{+/-}$ cross (1/4 of population is null and 3/4 heterozygous or homozygous wild-type) injected with hsfpq RNA, either wild-type (WT at low dose: 150 pg/embryo or high dose: 250 pg/embryo, to control for phenotype due to overexpression of normal proteins), or with N533H, L534I, or S660N mutations at

150pg/embryo. **B:** Lateral views of spinal cord motor neurons in embryos injected with SFPQ^{S660N} RNA at 48 hpf (Efthymiou and Gordon, 2020, data generated by Patricia Gordon).

We then questioned whether the mutated protein was normally localized in the embryonic neurons. We injected gfp-tagged human SFPQ or SFPQ^{S660N} RNA into the progeny of a zebrafish *sfpq*$^{+/-}$ incross at the one-cell stage and assessed localization using antibody staining at 36 hpf. We found that the hSFPQ^{S660N} protein is very robustly detected in all cell nuclei, similar to wild-type human SFPQ (Figure 31A-B). In addition, however, we detected large SFPQ^{S660N} puncta in a proportion of axons imaged (Figure 31A, 27C), substantially brighter and bigger than in the wild-type, indicating an abnormal organization of SFPQ protein complexes in these axons. The puncta were apparent in 4/19 embryos tested, all showing rescued null mutant morphology, indicating that these large SFPQ^{S660N} puncta only appear in an *sfpq*$^{-/-}$ background. Incross progeny injected with wild-type hSFPQ did not exhibit visible axonal puncta at the same magnification.

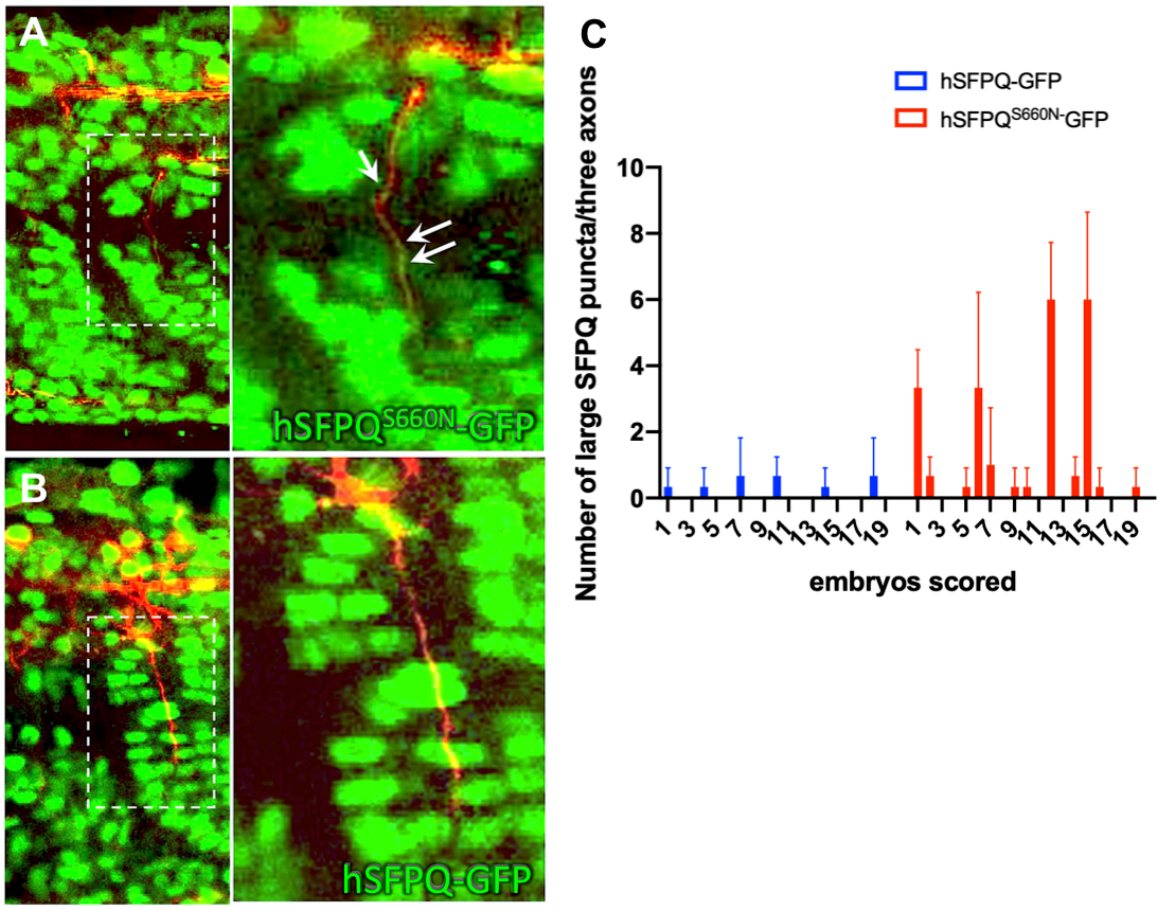


Figure 31. Distribution of human SFPQ protein in rescued zebrafish mutants. A, B: Lateral views of confocal imaging of trunk innervation at the level of somite 16 showing human SFPQ expression (red) in 36hpf *sfpq-/-* zebrafish. Axonal tracts stained with acetylated tubulin antibody (green). Arrows showing large puncta of protein in motor axons. **C**: Quantification of large puncta in three axons (somite 15, 16, 17) per embryo in 20 individuals at 36hpf, per rescue condition (wild-type or S660N) (Efthymiou and Gordon, 2020, data generated by Patricia Gordon).

3.3.3 Discussion

Taken together, these genetic and functional studies confirmed the vital role of the mRNA splicing factor *SFPQ* for nervous system development and function. They more importantly demonstrate the deleterious effect of the p.Ser660Asn missense mutation, affecting the ability of the SFPQ protein to maintain normal hindbrain and midbrain separation and

normal axonal growth and branching during embryonic development. The same function is likely to be conserved across species, thus supporting the conclusion of a set of deleterious neurological consequences of *SFPQ* biallelic p.Ser660Asn mutation in humans. Further cellular and model organism work remains to be done to fully understand the biological process associated with changes in SFPQ function, including skeletal muscle physiology or iron homeostasis, and to assess how biallelic or monoallelic *SFPQ* gene variants may cause or contribute to different neurological and degenerative features in humans.

Chapter 4. Neurodevelopmental disorders (NDDs)

4.1 Clinical and molecular aspects of neurodevelopmental disorders

Neurodevelopment is the biological process resulting in the development and maturation of the nervous system. In humans, the process starts at the third week of embryonic growth with the formation of the neural tube. From the ninth week onward, the brain maturates and acquires its typical structure, in a coordinated chain of events that include abundant cell proliferation, migration, and differentiation.[169-173] Any disruption to this process may lead to a dysfunctional brain development, and therefore to a neurodevelopmental phenotype.

Neurodevelopmental disorders (NDDs) are a group of complex conditions characterized by impairments in cognition, social communication, behaviour and/or motor skills resulting from abnormal brain development. Intellectual disability (also known as mental retardation), communication disorders, autism spectrum disorder (ASD), attention deficit/hyperactivity disorder (ADHD), impairment in vision and hearing, and schizophrenia fall under the umbrella of NDD. The prevalence of these disorders constitutes a serious health problem in modern days. They affect at least 1 in 10 children[174] and have an onset early in life, as early as during birth or the first months of life. While the symptoms of neurodevelopmental disabilities often change overtime or evolve as a child grows older, some disabilities can be permanent. Genetic screening, clinical diagnosis and treatment of these disorders can be difficult and treatment often involves a combination of professional therapy, pharmaceuticals, and home- and school-based programs.

Although the genetic aetiology of NDDs is far from being completely known, significant genetic advances have been made in the last years, ascertaining specific biological pathways underlying the molecular mechanisms of these disorders. The current mutational spectrum of NDDs includes many hundreds of genes related to neurodevelopmental

pathways such as those associated with chromatin remodelling, synaptic function, and transcriptional regulation. [175-177] There is considerable evidence for genetic heterogeneity within but also between different NDDs, with a significant overlap of genes involved in more than one NDD, and a constant increasing number of known causative genes.

Intellectual disability (ID) is another common neurodevelopmental manifestation that affects 1-3% of the world's population.[178] Genetic factors play a key role causing the congenital limitations in intellectual functioning and adaptive behaviour. The heterogeneity of ID makes it more challenging for genetic and clinical diagnosis, but the advent of large-scale genome sequencing projects in a trio approach has proven very effective. A combined approach of next-generation sequencing and functional, electrophysiological, and bioinformatics analysis has identified new ways to understand the causes of ID and help to interpret novel ID-causing genes. The expansion of genomic analysis of ID patients in diverse and ancient populations can reveal rare novel disease-causing genes such as the 700 genes already linked with AD, AR and X-linked forms of ID. [178]

However, most neurodevelopmental disorders have complex and multifactorial causes such as genetic, biological, psychosocial and environmental risk factors. Interestingly, a fraction of genetic risk factors for autism have been *de-novo* mutations. These types of mutations are not inherited from either of the parents' genes, but are mutations formed in parents' germ cells (either the egg or sperm) or potentially from exposure to contaminants (metals, air pollution, benzene and persistent organic pollutants) during pregnancy[179]. Many environmental contaminants have been identified as agents capable of causing mutations in DNA, by leading to oxidative DNA damage and by inhibiting the body's normal ability to repair DNA damage. Two gene variants identified as *de-novo* and linked to a neurodevelopmental phenotype during my PhD project will be described further in this thesis (*ASXL1* and *KAT6A*).

WES, especially when done in family trios, has played an important role in neurodevelopmental disorders, because it has allowed us to identify both *de-novo* mutations and inherited risk variants with variable penetrance. The next part of this chapter will give an in-depth description of one such study carried out in our lab by myself and Dr. Andreea Manole (postdoctoral researcher) and Dr. Emer O'Connor (neurologist), leading to the identification of *NARS1* as a candidate gene for a complex form of early-onset neurodevelopmental delay accompanied by epilepsy, neuropathy and developmental deformities.

Various other collaborative projects were carried out during my PhD project involving patients recruited from the SYNaPS patient consortium. Salpietro *et al.* (2019) reported *de-novo* mutations and microdeletions involving *GRIA2* as a cause of NDDs and developmental epileptic encephalopathy and underlined the importance of the GluA2 subunit in the regulation of Ca^{2+} permeation and voltage rectification of AMPARs and therefore in human synaptic plasticity and brain development and function.

Dias *et al.* (2019) reported on the genotype-to-phenotype relationship with *NTNG2* variants, and established an initial description of the clinical spectrum of patients carrying rare biallelic *NTNG2* variants. This gene should be considered in the clinical evaluation of children with severe intellectual disability and neuropsychiatric symptoms (autism) and motor impairment, and very importantly it should be added to clinical gene-panel tests for intellectual disability given the marked yet variable clinical phenotype.

4.2 A *de-novo* truncating mutation in *ASXL1* linked to segmental overgrowth

4.2.1 Introduction

Bohring-Opitz syndrome (BOS) is a rare autosomal dominant (AD) syndrome, characterized by the variable combination of severe intrauterine growth restriction, feeding difficulties, hypotonia, profound intellectual disability (ID), trigonocephaly with distinctive facial features, peculiar facial *nevus flammeus* and broad MRI abnormalities (e.g., hypomyelination, hypoplastic corpus callosum, spinal cord abnormalities). BOS has been reported in literature to be caused in up to 75% of cases by *de-novo* truncating variants in *ASXL1*. This gene encodes the additional sex combs-like 1 protein involved in chromatin remodelling. These epigenetic processes are mediated either by histone H2A deubiquitination 1 and/or transcriptional regulation of the polycomb group repressor complex 2 mediated homeobox (HOX) gene. HOX genes are critical for body patterning and segmental identity during human embryogenesis.[180] Bone marrow stromal cells (BMSCs) are an example of such homeotic genes that need to be regulated by ASXL1 to achieve a homeostatic balance between self-renewal and differentiation. An imbalance in the cell fate or cell lineage commitment of BMSCs can lead to developmental or skeletal defects, as observed in BOS cases.[181] The ASXL1 protein is composed of an N-terminal DNA binding domain made up of the ASXN and ASXH sub domains, the ASXM1 and ASXM2 domains in the middle region as well as a PHD (zinc finger) domain in the C-terminal region, specifically important for mediating interactions with other proteins (Figure 32D). *ASXL1* mutations have been identified as a cause of BOS in 2011[180] and most individuals reported so far have been found to carry *de-novo* truncating mutations. (Table 19) [182-187]

Author	Year	Cases	Mutations
Urreizti *et al.*	2018	1 case	p.G646W*fs**12
Bedoukian *et al.*	2018	1 case	**p.R965***
Carlston *et al.*	2017	1 case	**p.R404***
Arunachal *et al.*	2016	1 case	p.G680R*fs**38
Dangiolo *et al.*	2015	1 case	p.F1373*fs*
Russell *et al.*	2015	8 cases	p.C672W*fs**4, p.G1026fs, p.G642*, p.L775*, p.I1919fs, p.Y425Q*fs**12, **p.S846*fs**5**
Urreizti *et al.*	2015	1 case	p.P701S*fs**16
Magini *et al.*	2012	2 cases	p.E803T*fs**17, **p.R965***
Hoischen *et al.*	2011	7 cases	**p. R404*,** p.Q733*, p.Q778*, p.L823*, **p.S846Q*fs**5,** p.Q925*, p.S1028*

Table 19. Previously reported genetically confirmed *ASXL1* cases with BOS. The three most common mutations are shown in bold.

Mutations in genes involved in chromatin remodeling have been implicated in broad phenotypes of congenital abnormalities and neurodevelopment. However, limited genotype–phenotype correlations are available for some of the rarest genetic disorders that affect chromatin regulation. This work hereby describes a 12-year-old girl presented at birth with severe hypotonia, developmental delay, a mid-line capillary malformation and distinctive craniofacial features. During the natural history of her disease the girl developed severe spasticity and drug-resistant seizures, leading to a diagnosis of Bohring-Opitz syndrome (BOS). We performed whole exome sequencing (WES) and identified a *de-novo* mutation in *ASXL1* (c.2033dupG) which results in the introduction of a premature stop codon (p.R678fs*6). *ASXL1* encodes a polycomb repressive complex protein implicated in chromatin regulation and de-novo mutations are a known cause of BOS. Phenotypes with segmental craniofacial overgrowth associated to midline capillary malformations enlarge the clinical spectrum of BOS at onset and further expand the differential diagnosis in *ASXL1* mutation carriers.[188]

4.2.2 Results

Clinical findings

Here, we report on the clinical characterization of one young girl affected with atypical BOS, whose phenotypic manifestations included global developmental delay with absent speech and severe motor impairment. The clinical data was provided by Dr. Gabriella di Rosa. She presented with failure to thrive, hirsutism, glabellar *nevus flammeus* (observed since birth) and trigonocephaly with distinctive craniofacial features including prominent globes, widely set eyes, synophrys (fusion of eyebrows), and micrognathia.

In previous generations neurological or genetic diseases were absent in the family and parents were healthy (Figure 32A). Pregnancy was complicated by threatened abortion and intrauterine growth restriction documented by repeated ultrasound scans. The girl was born at 38 weeks with weight <10th centile for gestational age. At birth, generalized low muscle tone as well as small hands with contractures of the fingers and a *nevus flammeus* on the face were observed. On the third day of life she suffered from an epileptic seizure characterized by peri-oral cyanosis, hypertonia and eye deviation. Axial hypotonia was observed since the neonatal age and a severe delay of motor milestones was noticed in the first months of life. During the first year of life she started to experience frequent (almost daily) episodes of generalized hypertonic seizures with lateral eye and mouth deviation, eyelid flutter and oral automatisms. Seizures were usually followed by hypotonia and regression of motor developmental milestones. Poor visual interaction was noticed since the first months of life as well as myopia later on in life. Severe language impairment was also present, and the girl never attained intelligible speech. She presented with microcephaly (OFC <10th) and trigonocephaly as well as distinctive facial features including prominent eyes, micro and retrognathia.

On her neurological examination at the age of 15 years old, she presented with severe spastic paraparesis and dystonic posturing of the hands, wrists, elbows as well as foot and brisk deep tendon reflexes. Her shoulders were externally rotated and adducted, elbows and wrists flexed in ulnar deviation, and ulnar deviation of the metacarpophalangeal joints. The patient was treated during her first year of life with antiepileptic medication, specifically sodium valproate, and since the second year of life with levetiracetam, with good clinical response, as she was seizure-free until the age of 8 years. Extensive diagnostic and metabolic work-up was reported to be normal. Auditory brain response showed sensorineural hearing loss only in the left ear. The patient also underwent also molecular investigation such as karyotype and array comparative genome hybridization (array-CGH) that were both reported as normal.

At the age of 5 years, electroencephalogram (EEG) activity recorded during sleep showed recurrent interictal high-voltage, bilateral, spike-wave discharges, followed by brief sequences of slow delta activities, interspersed with brief tracts of diffuse background slowing according to a fragmented hypsarrythmia pattern. A progressive organization of background rhythm occurred. At 8 years of age, EEG activity was characterized by interictal high-voltage, bilateral, spike-wave discharges, at 3 Hz, on bilateral post-central regions. Brain magnetic resonance imaging (MRI) performed at the age of 2 years, revealed brain asymmetry with reduced white matter volume in the left hemisphere, associated with enlargement of the left lateral ventricle and widening of the left hemispheric sulci; the corpus callosum is slightly thinned. These MRI features have been reported in BOS before.[185] The asymmetric appearance of the brain parenchyma appeared of mild degree, cortical abnormalities such as polymicrogyria, frequently associated with this syndrome, hydrocephalus or thickened corpus callosum were also noticed. [189]

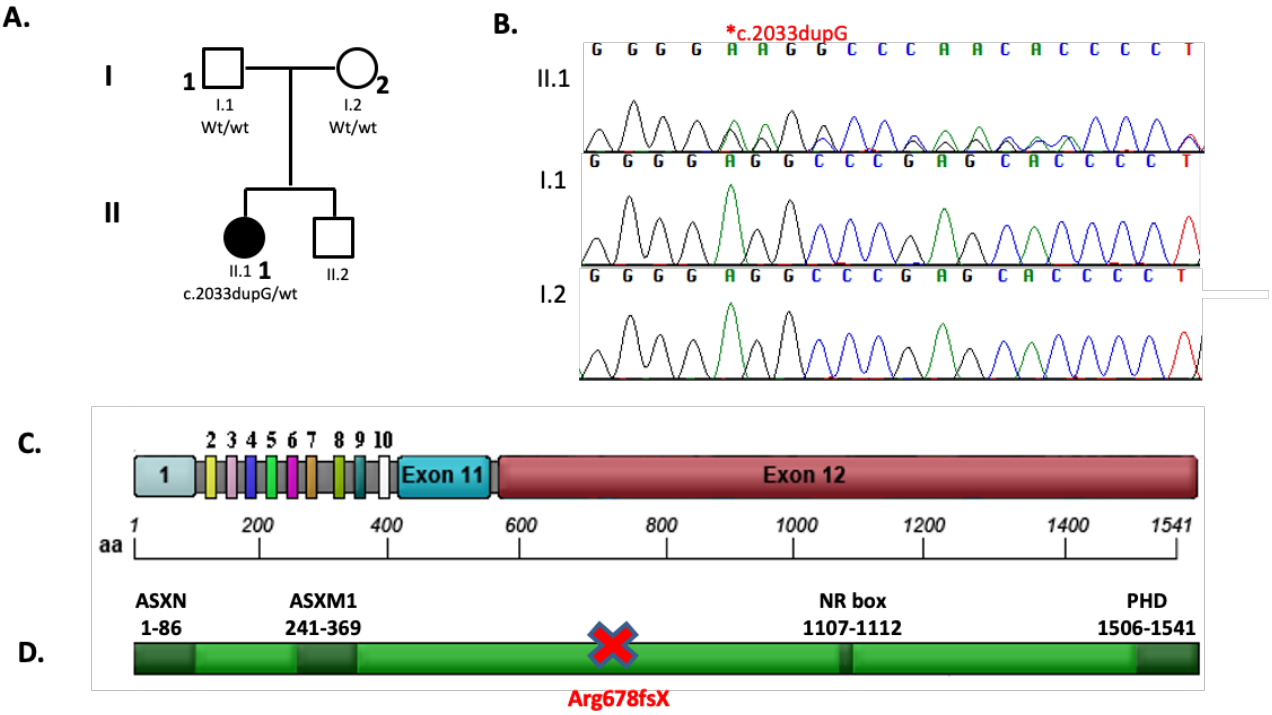


Figure 32. Pedigree and Sanger sequencing of the ASXL1 patient. **A.** The pedigree diagram of the family with an *ASXL1* mutation. **B.** Individual results of Sanger sequencing showing the proband (II.1) carrying the heterozygous *ASXL1* truncating mutation (c.2033dupG, p.R678fs), while being absent in the two parents I.1 and I.2 **C.** Representation of *ASXL1* gene showing the amino acid count. **D.** Representation of the protein domains: ASXN, conserved domain at the N-terminus; ASXM, conserved domain in the middle part; NR, nuclear receptor; PHD, plant homeodomain, and the position of the frameshift mutation in *ASXL1* (c.2033dupG, p.R678fs) as indicated (Efthymiou *et al.,* 2019).

Genetic findings

In total, 55,531,100 (II-1) unique reads were generated. Sequencing data was reviewed for evidence of somatic mosaicism in both the proband and the parents' samples as the *de-novo* variant was only present in 16% of exome sequencing reads and present at an allele ratio of < 35% in the proband. We used a filtering strategy as described before (General methods) fitting a recessive or a *de-novo* model.

We have not identified rare potentially damaging recessive variants in the girl, and a *de-novo* truncating mutation in *ASXL1* (c.2033dupG), resulting in the introduction of a premature stop codon six amino acids downstream (p.R678fs*6), emerged as the most likely explanation for the disease pathogenesis. This is supported by the severe impact of the *ASXL1* mutation on protein function and the existing reports linking loss of function (LoF) variants in this gene to a similar clinical phenotype. [180; 182-187] In addition, we screened *ASXL1* in >1,000 trios exomes from individuals affected with genetically undiagnosed neurodevelopmental diseases recruited within the SYNAPS Study Group research initiative (neurogenetics.co.uk/synaptopathies-synaps), without identifying any additional *de-novo* variant in this gene. Segregation analysis by traditional Sanger sequencing [primers used: forward (5'-GGTCACCACTGCCATAGAGA-3') and reverse (5'-GAGGATAAGGCGGCAGTAGT-3')] confirmed that mutation occurred *de-novo* in the girl (Figure 32B). The patient has provided informed consent for publication of the case.

Most *ASXL1* mutations reported so far have been found in the last exon (exon 12) of the gene, that includes ~50% of the whole coding region of the gene, and where truncating variants are likely to generate transcripts that can escape nonsense mediated decay. The mutation that we identified occurs in exon 12, leading to a premature truncation of the downstream carboxyterminal plant homeodomain finger. Thus, the mutation may affect the binding of ASXL1 to chromatin, altering the activation or silencing of transcription factor genes involved in embryonic development (Figure 32D). During the natural history of the disease in the girl, additional progressive craniofacial and neurological features became evident, leading to a clinical diagnosis of atypical BOS.

4.2.3 Discussion

Previous work in mice and Drosophila showed significant variability with the phenotypes observed in BOS individuals carrying *ASXL1* truncating mutations. *ASXL1*$^{-/-}$ mice present multiple developmental abnormalities, including anophthalmia, 30% reduction in body and skull size, resembling the microcephaly and changes in cranial shape observed in BOS individuals, as well as cleft palate and mandibular malformations.[181] This reinforces the role of ASXL1 as an essential protein needed for multi-lineage differentiation potential such as osteogenesis, adipogenesis and further skeletal and body development and potentially explain both the skeletal and brain asymmetries that we observed. Our patient was initially evaluated for the combination of her facial *nevus flammeus*, abnormal facial asymmetry (left > right), failure to thrive and global developmental delay and at that point a segmental overgrowth disorder associated to midline capillary malformation was also suspected as part of the differential diagnosis for her condition.

Segmental overgrowth anomalies associated with midline capillary malformation enlarge the clinical spectrum of BOS at onset and further expand the differential diagnosis in *ASXL1* mutation carriers, highlighting a possible involvement of chromatin remodeling and epigenetic regulation in the developmental and skeletal defects associated with BOS.

Mutations in other chromatin remodeling genes such as DNMT3A have been linked with developmental disorders associated to segmental overgrowth disorders, such as PIK3CA related phenotypes. Future studies are needed to fully assess potential epigenetic changes due to mutations in the above genes and their impact on molecular pathways associated with segmental overgrowth disorders associated with capillary anomalies.

4.3 A loss-of-function homozygous mutation in DDX59 implicates a conserved DEAD-box RNA helicase in nervous system development and function

4.3.1 Introduction

Postaxial polydactyly (PAP) is the occurrence of a supernumerary sixth digit of the hand and/or the feet, and can be observed in the context of several clinically and molecularly heterogeneous genetic disorders.[190] The occurrence of PAP and intellectual disability (ID) is often due to mutations in genes involved in ciliogenesis (e.g., *BBS2*, *WDPCP*, *KIAA0586*, *TCTN1*, *TCTN2*, *MKS1*, *TMEM67*, *CC2D2A*) and most of the times these phenotypes are associated with a wide array of multiple, variable (e.g., skeletal, ophthalmological, hepatic, renal, and genitourinary) abnormalities.[191; 192] However, the combination of autosomal recessive PAP and neurological involvement in the absence of other symptoms is very rare, with few families reported in the literature as Oliver syndrome (OS; MIM #258200), a rare distinct clinical phenotype with no causative gene yet identified. [193-195]

This chapter reports on a homozygous frameshift deletion in DDX59 (c.185delT; p.Phe62fs*13) in a family presenting with postaxial polydactyly and a neurological phenotype characterized by microcephaly, intellectual disability, epilepsy, and white matter signal abnormalities associated with cortical and sub-cortical ischemic events. DDX59 encodes a DEAD-box RNA helicase and its role in brain function and neurological diseases is unclear. We showed a reduction of mutant cDNA from patient-derived cell lines; furthermore, analysis of human brain gene expression provides evidence that DDX59 is enriched in oligodendrocytes and might act within pathways of leukoencephalopathies associated genes. In collaboration with Dr. Mutsuddi (Varanasi,

India), we also characterized the neuronal phenotype of the Drosophila model using mutant mahe, the homologue of human DDX59 previously shown to regulate Notch signaling in fruit-flies and showed that mahe loss-of-function mutant embryos exhibit impaired development of peripheral and central nervous system. Taken together, our results support a conserved role of this DEAD-box RNA helicase in neurological function.[196]

4.3.2 Results

Clinical findings

Using a WES approach described in Chapter 2, we investigated an Italian family previously reported as having OS.[195] The clinical data was provided by Dr. Maria Cutrupi and Dr. Valeria Dipasquale. The two probands presented a typical orofaciodigital syndrome phenotype with PAP, subtle midline anomalies (i.e., cleft lip) and distinctive facial features. In addition, they also had a heterogeneous neurological involvement which included delay of developmental milestones, ID, infantile-onset seizures, lower limbs weakness and neuropathy, and adult-onset white matter signal abnormalities associated to episodes of ischemic strokes. There was no history of previous neurological or genetic diseases in the family and the pedigree suggested autosomal recessive inheritance (Figure 33A).

Patient 1 (Figure 33A, II-1) was the first born from healthy parents, non-consanguineous (as far as they were aware). Family history was unremarkable, except for three prior spontaneous miscarriages. The pregnancy was complicated by intrauterine growth retardation (IUGR). Delivery at term was normal, with a weight at birth of 2,350 g (<3rd centile), length of 47 cm (3rd centile), and occipital-frontal circumference of 32 cm (5th centile). APGAR scores were 6 and 9 at 1 and 5 min, respectively. He had bilateral postaxial extra-digits on his hands that were surgically removed in his late childhood. He

also had bilateral cutaneous syndactyly of fingers 2–5, clinodactyly of the fifth fingers, and fingertip pads. His lower limbs were normal. Since the first months of life, he developed generalized seizures, which were controlled by anticonvulsant drugs. Developmental milestones were delayed and the patient showed cognitive difficulties during childhood, with an I.Q. of 70 (Terman-Merril scale) measured at the age of 9 years. For these reasons, he has undergone physiotherapy, occupational and speech therapy since the age of 3 years. At the age of 17 years his height was 165 cm (3rd centile), weight was 67 kg (50th centile), and head circumference 53 cm (10th centile). He had distinctive facial features, including prominent thick eyebrows, malocclusion, high-arched palate, and rounded prominent jaw.

A cerebral magnetic resonance imaging (MRI) disclosed thinning of the cerebral cortex in front of the ventricular collateral trigone (not shown). Wakefulness electroencephalograms (EEG) showed diffuse high amplitude slow waves intermingled with sharp waves or spikes. Since early adulthood he complained of migraine headaches. As part of his neurological presentation, he also presented lower limb weakness and some walking difficulties. At the age of 30, after an episode of loss of consciousness associated with a generalized seizure, he underwent a follow-up brain MRI scan which showed diffuse white matter signal abnormalities and multifocal cortical-subcortical infarcts involving both cerebral hemispheres. Extensive metabolic and genetic investigations, which included array comparative genome hybridization (array-CGH) and panel sequencing for 22 leukoencephalopathies-associated genes, were performed and fully reported as normal.

Patient 2 was the younger sister of Patient 1 (Figure 33A, II-2). Pregnancy was uncomplicated, and delivery at term was normal. Her birth weight was 2,850 g (10th centile), length was 47 cm (3rd centile), and occipital-frontal circumference was 33 cm (10th centile). Her APGAR scores were 8 and 9, at 1 and 5 min, respectively. From 4

months of age, she developed generalized tonic-clonic seizures, which were controlled by anticonvulsant drugs. She had postaxial polydactyly of the left hand, with a camptodactylous extra digit, bilateral clinodactyly of the fifth fingers, cutaneous syndactyly of fingers 2–5, and prominent fingertip pads. Postaxial polydactyly was also present on the right foot, with bilateral brachydactyly of toes 3–5. She also had malocclusion, high-arched palate, and thoracic right convex lateral scoliosis. Similarly, to her brother, psychomotor development was delayed, with an I.Q. of 68 (Terman-Merril scale) at the age of 7 years. She also had speech difficulties. At the age of 13 years, her height was 137 cm (3rd centile), weight was 34 kg (10th centile), and occipital-frontal circumference was 50 cm (<3rd centile). Extensive laboratory tests, including metabolic studies, karyotype, array-CGH and FRAXA analyses were normal. A cerebral MRI showed thinning of the cerebral cortex in front of the ventricular collateral trigone. A follow-up MRI performed at the age of 21 years showed signal abnormalities in the sub-cortical and deep white matter of both cerebral hemispheres, with more focal cortical-subcortical gliosis in the right frontal lobe. She presented with weakness of the lower limbs and motor nerve conduction studies showed a mild reduction of motor conduction velocities with peroneal amplitude of 2.4 mV, and conduction velocity of 42.9 m/s [low limits (3rd per age and height): amplitude (mV) 2.6, conduction velocity (m/s) 43], suggesting a mild axonal neuropathy. The patients have provided informed consent for publication of the cases.

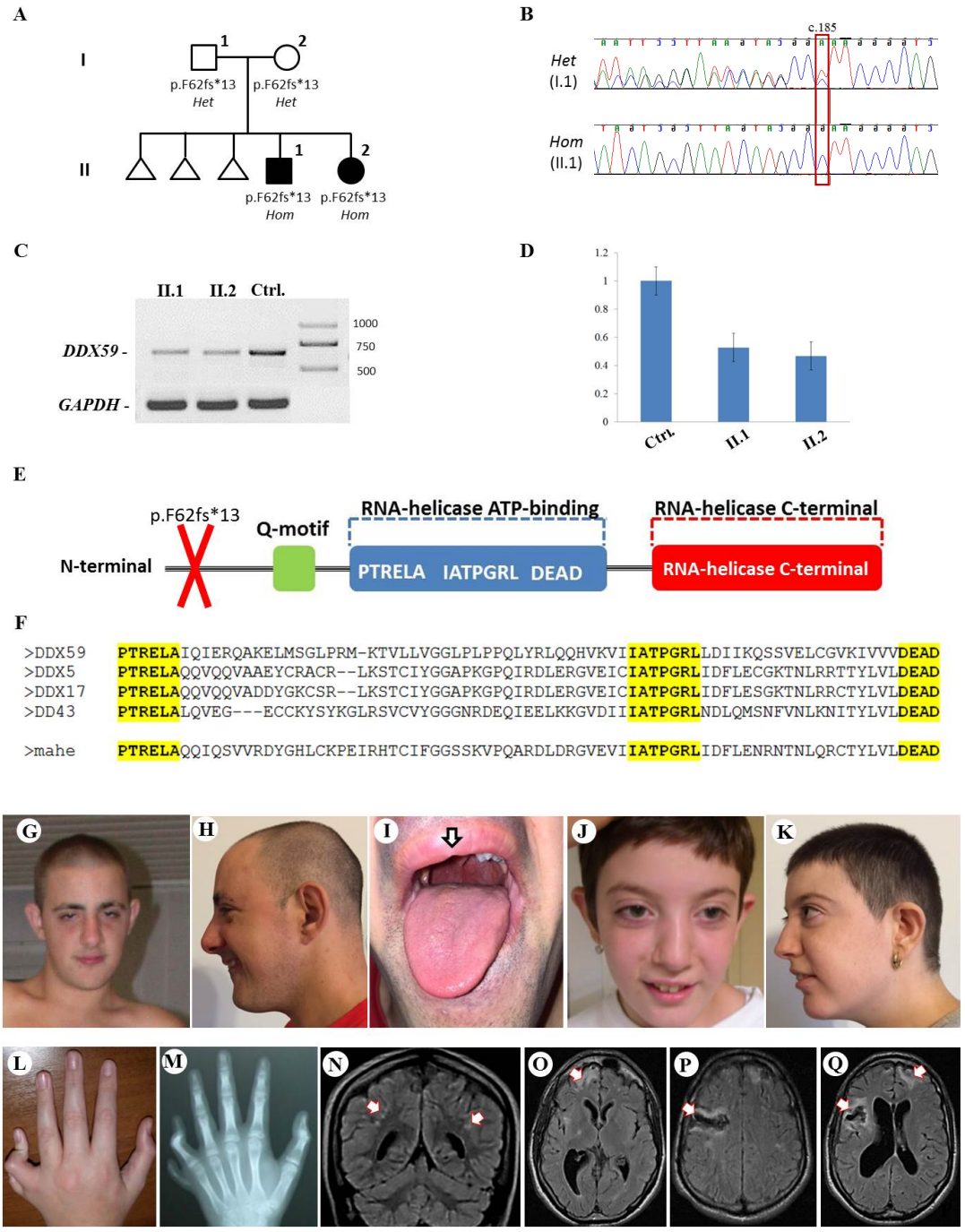


Figure 33. Family tree, Sanger sequencing, radiological scans and *DDX59* mutation analysis. (A) Pedigree of the Family. (B) Electropherograms of 1 carrier parent (I.1) and 1 index case in reverse orientation (II.1) with the heterozygous (Het) and homozygous (Hom) c.185delT *DDX59* variant, respectively. (C) Semi quantitative polymerase chain reaction (semi-qPCR) amplifying the mutant cDNA transcript from mRNA extracted from

the immortalized lymphoblastoid cell lines of the two affected siblings and a wild-type (age-matched) control (CTRL). (D) Analysis of the semi-qPCR using the densitometry software ImageJ after normalization relative to a housekeeping gene (GAPDH), and calculation using a relative relationship method. (E) Schematic representation of the Ddx59 protein. (F) Multiple-sequence alignment showing complete conservation of DED-BOX RNA Helicase active domains (PTRELA, TPGR, DEAD) sequence across DDX59, Mahe, and the other Mahe homologues (*DDX5*, *DDX17*, *DDX43*) (G) Patient II.1 at the age of 19 years, not the prominent, thick eyebrows, malocclusion, high-arched palate, and rounded and prominent jaw. (H) Patient II.1 at the age of 29 years. (I) Patient II.1, note the subtle midline defect. (J) Patient II.2 at the age of 13 years, note the distinctive facial features similar to her elder brother. (K) Patient II.2 at the age of 24 years. (L) Left hand postaxial polydactyly and camptodactyly of Patient II.2 at the age of 13 years. (M) Skeletal X-ray of the hands of Patient II.2 at the age of 13 years. (N) Brain MRI of Patient II.2, note in the coronal scan the diffuse white matter hyperintensities. (O, P, and Q) Magnetic resonance imaging (MRI) of the brain of Patient II.1 at the age of 27 after a stroke-like episode; note in the axial scans the diffuse subcortical infarction in the right hemisphere mainly involving frontoparietal lobes; also note the diffuse white matter hyperintensities (modified from Salpietro and Efthymiou, 2018).

Genetic findings

In total, 83,572,847 (II-1) and 81,527,162 (II-2) unique reads were generated from WES. After applying filtering criteria (see chapter 2), no plausible shared compound heterozygous variants were identified by WES; there were however 3 genes carrying rare (likely) damaging variants, according to guidelines for variants interpretation[71], which were homozygous in both probands (Table 20). Two out of these 3 three variants were missense changes not consistent with the phenotype and also identified in additional (non-affected) individuals from our in-house exome database, containing over 4,000 exomes.

Table 20. Homozygous likely damaging variants identified in the two affected siblings by whole-exome sequencing (WES).

PKN1	(NM_213560.1)	c.530 A>G; p.(Gln177Arg)
ALPPL2	(NM_031313.2)	c.779 A>G; p.(His260Arg)
DDX59	(NM_001031725.4)	c.185delT; p.(Phe62fs*13)

A homozygous frameshift deletion in DDX59 (NM_001031725.4; c.185delT: p.Phe62fs*13) emerged as the most likely explanation for the disease pathogenesis; this is also supported by a more severe impact of the mutation on protein function (truncating vs. missense) and an existing report previously linking this gene to a similar (albeit milder) phenotype of Orofaciodigital syndrome type V (OFD5, MIM #174300) with PAP and ID.[197] Also, that study shows expression patterns and indicates a likely important role of this gene in midline development and the nervous system.[197] Segregation analysis performed by traditional Sanger sequencing confirmed the mutation homozygous in the two affected siblings and heterozygous in both their parents. The identified *DDX59* homozygous variant (NM_001031725.4; c.185delT: p.Phe62fs*13) was submitted to the Leiden Open Variation Database (www.lovd.nl/; variant ID #0000221973).

Functional investigation

The semi-qPCR from patients and an age-matched control's lymphoblastoid cell lines did not show complete mRNA degradation via NMD, although there was a reduction of the mutant cDNA transcript compared to wild-type controls (Figure 33C, D).

DDX59 encodes for a putative RNA helicase belonging to the DEAD-box family of proteins. These proteins are involved in various aspects of RNA metabolism through the evolutionary conserved ATP-binding domain, a core element which contains active motifs (PTRELA, TPGR and DEAD) required for the helicase activity.[198; 199] The exact role in nervous system and neurological disease of DDX59 is unknown. The gene has 8 exons, its transcript (ENST00000331314.6) contains 2289 nucleotides and the encoded protein is

619 amino acids long. The c.185delT results in a frameshift at amino-acid residue 62, generating a premature stop codon 13 amino acids downstream and a truncated protein with the helicase ATP- binding and C-terminal domains (and the evolutionary conserved active motifs PTRELA, TPGR and DEAD) being omitted (Figure 33E, F).

Homozygous mutations in *DDX59* were reported in only two studies so far, with 2 Arab and 2 Pakistani families presenting an orofaciodigital syndrome phenotype (with distinctive facial features and digital and midline abnormalities) associated to ID.[197; 200] The functional analysis in the original study from Shamseldin *et al*. (2013) showed that Dead-Box 59 RNA Helicase is involved in the ciliary Sonic Hedgehog Homolog (SHH) signaling, a pathway known to be implicated in ciliogenesis.[201]

The analysis of
Early Infantile Epileptic Encephalopathy
results showed a slight increase of the SHH protein in the patients (more significant in the eldest sibling) compared to the age-matched control (Figure 34), suggesting a possible downstream perturbation of the Hedgehog pathway due to loss-of-function mutation in *DDX59*.

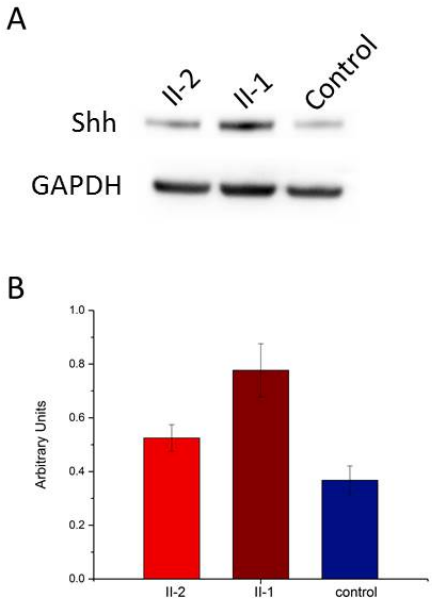


Figure 34. Western blot analysis of the SHH protein in patients and control derived lymphoblastoid cell lines show perturbation of SHH (and slight increase of the SHH protein) in *DDX59* mutation carrier (Salpietro and Efthymiou, 2018).

To better understand the white matter and sub-cortical changes as part of the neurological phenotype of this family, we analyzed *DDX59* expression in the human central nervous system (CNS) using genome wide transcriptomic data generated from control post-mortem human brain tissues as previously described.[202] In humans DDX59 is known to be expressed in the developing and adult human brain across all brain regions (Figure 35). Given that individuals with homozygous loss-of-function mutations in *DDX59* have an unusual neurological phenotype characterised by adult onset subcortical ischemic and white matter signal changes, we *in-silico* analyzed the human brain transcriptome of *DDX59* and noted that the gene is enriched in human intra-lobular white matter and then focused further analyses on this brain region. In particular, we used weighted gene co-expression analysis (WGCNA) to identify a module of genes, which are highly co-expressed with DDX59 in white matter.

This analysis demonstrated that DDX59 is located within a gene co-expression module, containing 1303 genes, with a module membership of 0.63 indicating confident assignment of DDX59. Interestingly, the white matter co-expression module containing DDX59 was significantly enriched for oligodendrocyte markers (p-value = 7.20×10^{-12}). Given the significant vascular and white matter changes evident from brain MRI scans for patients with homozygous loss-of-function variants in *DDX59*, we also investigated the DDX59-containing white matter module for the presence of genes already implicated in hereditary cerebral small vessel disease and stroke, or leukoencephalopathy. This demonstrated significant enrichment of genes associated with leukencephalopathy (p-value = 0.033), but no significant enrichment of genes associated with genetic forms of cerebral small vessel

disease (Figure 36). Thus, our analysis suggests that individuals with homozygous LoF variants in *DDX59* have a form of leukencephalopathy.

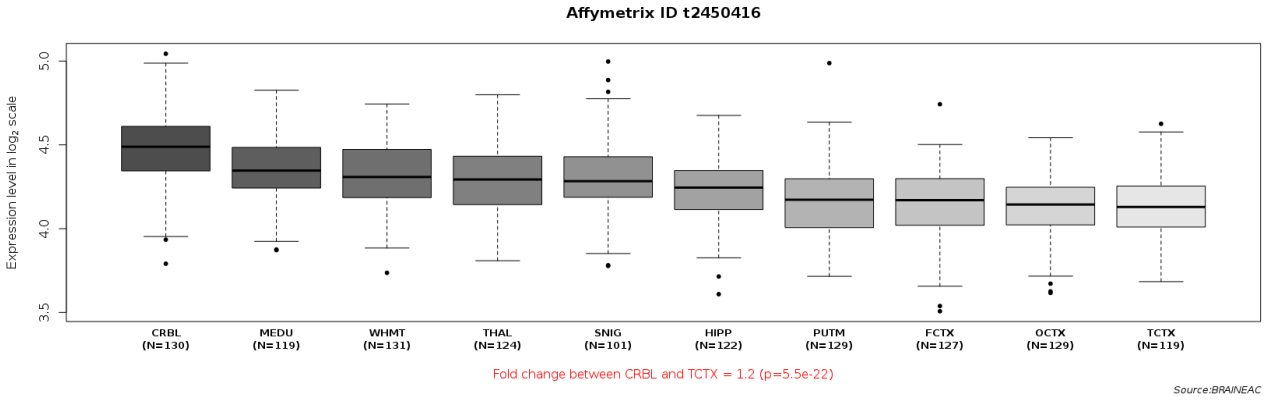


Figure 35. Expression for *DDX59* using transcriptome analysis by Affymetrix, showing diffuse expression of the gene through different the brain regions including cerebellar cortex (CRBL), medulla (MEDU), white matter (WHMT), thalamus (THAL, substantia nigra (SNIG), hippocampus (HIPP), putamen (PUTM), frontal cortex (FCTX), occipital cortex(OCTX) and temporal cortex (TCTX).

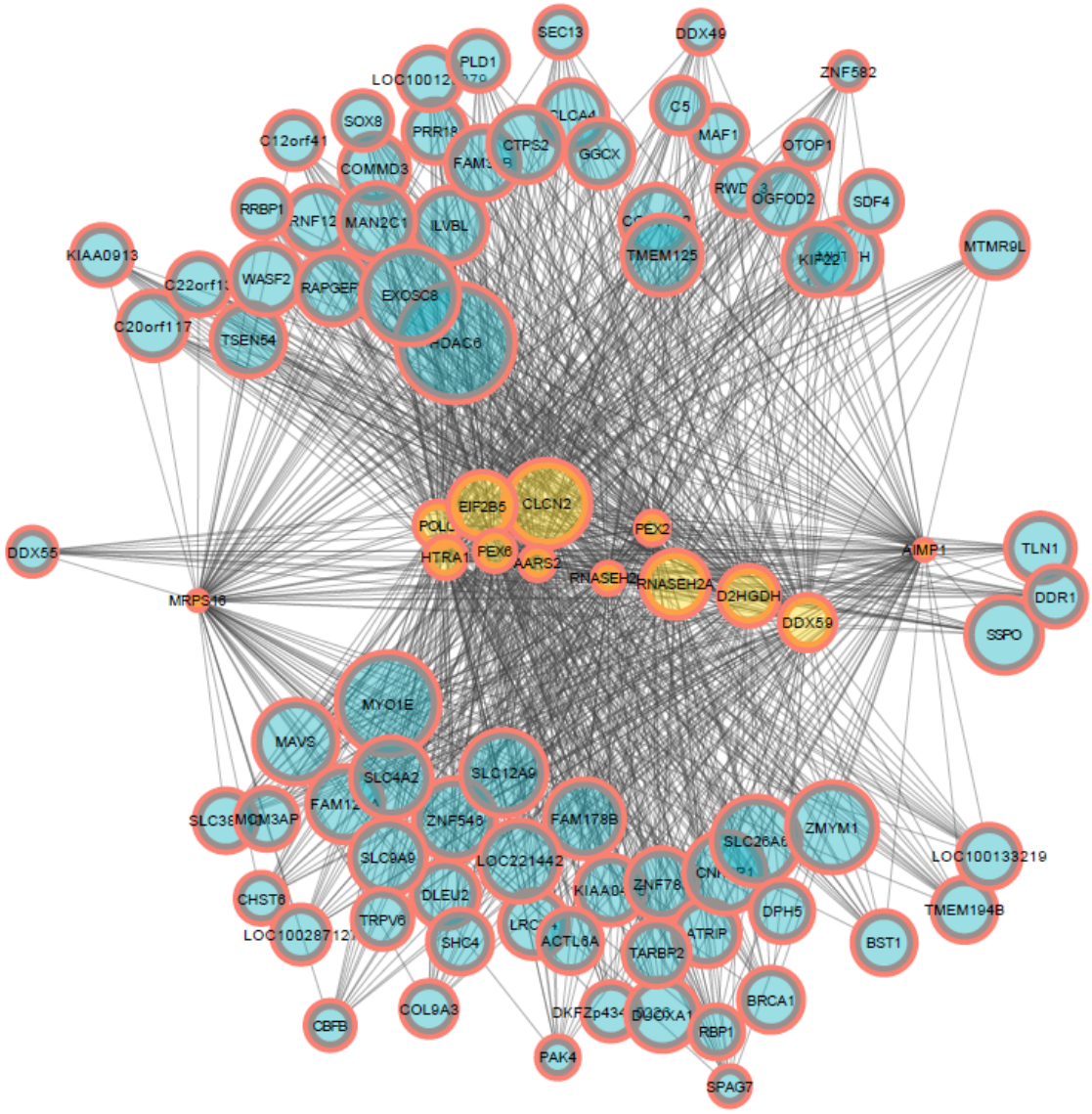


Figure 36. Bottom-up plot for *DDX59* and the leukoencephalopathy genes used as seed genes (highlighted in yellow). The algorithm adds, sequentially, the most connected gene to each seed gene (up to seven context genes per seed gene). Note all genes belong to the salmon module (color of the outer ring for each blob). In terms of connectivity to the context genes, DDX59 is allocated together with D2HGDH (data generated by Mina Ryten).

Significant defects in both peripheral nervous system as well as in the developing ventral nerve cord of the mahe (the homolog of human DDX59) mutant embryos (Stage-15) were seen in comparison to those of wild-type embryos. In addition, we carried out a survival

assay using homozygous hypomorphic allele maheEP1347, and observed shortened life span of viable mutant flies in comparison to wild-type flies (Figure 37).

Ciliopathies include several partially overlapping syndromes (e.g. Joubert syndromes, Meckel–Gruber syndrome, Bardet–Biedl syndromes) all characterized by pronounced neurodevelopmental features with significant abnormalities in central nervous system[203]. Paralleling the patient's symptoms harbouring mutation in DDX59, loss-of-function mutants of mahe also shows strong perturbation of the developing CNS, where a massive disorganization of the midline longitudinal axons were observed (Figure 37). Severe loss or incomplete ventral nerve cord along with gaps observed in mahe null embryos reflects similar brain malformation phenotype like that of ciliopathy associated syndromes. Of note, earlier studies have reported that in Bardet-Biedl syndrome in which mutation in genes associated with ciliopathy were identified showed pathological phenotype in the peripheral neurons[204]; thus, defects in peripheral nervous system observed in mahe mutant embryos may be because of a similar outcome. Unlike mahe mutants which have isogenized genetic background, human patients may have additional variants in genes which could modify the peripheral neurological phenotype associated with biallielic loss of DDX59.

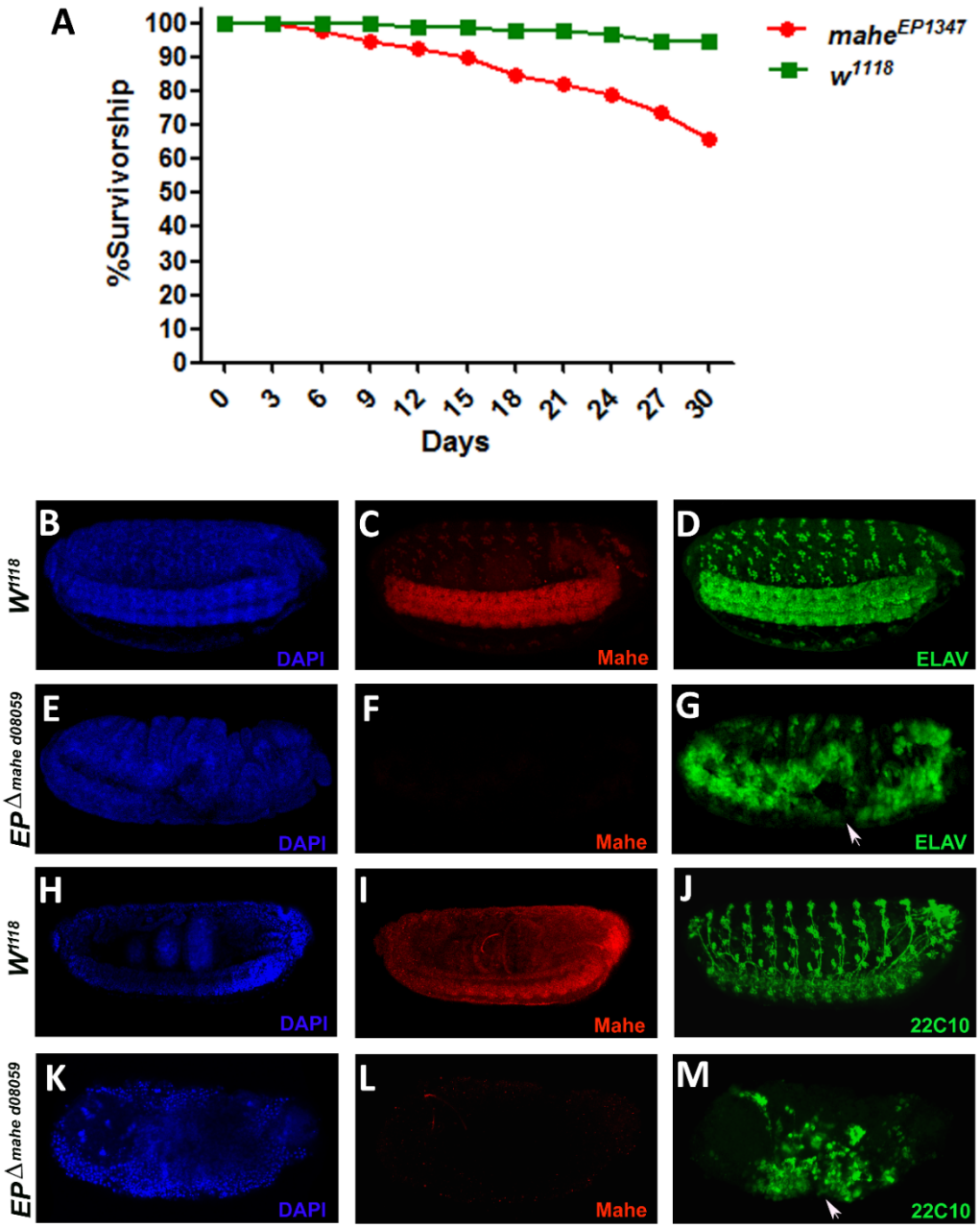


Figure 37. Neurological features of Mahe loss-of-function embryos. Mutations in *mahe*, the *Drosophila* homologue of *DDX59* reduces lifespan and also displays neuronal defects(A) *mahe*^{EP1347} hypomorphic alleles were assayed for lifespan, graph represents shortened life span of mutants in comparison to that of *w*^{1118} (wild-type). (B-M) Stage 15 embryos immunostained with anti-ELAV and 22C10, both of which are neuronal markers, were used to visualize the developing nervous system and DAPI was used to mark the nucleus. (B, C, D and H, I, J) Wild-type embryos showing normal peripheral nervous system and ventral nerve cord development along with *mahe*, expression. (E, F, G and K, L, M) *EP*$^{\Delta mahe\ d08059}$ (*mahe* null) embryos showing absence of Mahe protein and neuronal

markers revealed severe defects (arrows) like gap in ventral nerve cord and gross disorganization in peripheral nervous system during embryonic nervous system development (Salpietro and Efthymiou, 2018, data generated by Mousumi Mutsuddi).

4.3.3 Discussion

Of importance, RNA helicases from the DEAD-box family are found in almost all organisms and have crucial roles in many aspects of RNA metabolism, including transcription, pre-mRNA splicing, translation initiation, RNA transport and decay.[198; 199] However, despite their importance in cellular processes and involvement in several molecular pathways, only relatively few examples of mutations in DEAD-box RNA helicase genes have been reported so far in association to monogenic human diseases and none of them with biallelic truncating mutations.[197; 205]

In our family, the two patients presented autosomal recessive postaxial polydactyly associated with a complex neurological involvement. They presented with small OFC since birth and we propose this could be regarded as a consistent feature of the DDX59-related syndrome, given the occurrence of congenital microcephaly also in the patients recently reported.[200] They had also global neurodevelopmental delay and episodes of generalized seizures with onset in the first months of life. Patient II.1 presented since early adulthood diffuse white matter signal abnormalities associated with sub-cortical ischemic changes, Patient II.2 showed brain MRI features similar to her brother and EMG features of mild peripheral (axonal) neuropathy. Notably, the adult-onset abnormal findings on brain imaging observed in the two probands resemble those reported in CADASIL syndrome, caused by mutations in the NOTCH3 gene, although the white matter involvement in the latter consist in a more confluent leukoencephalopathy pattern, which is associated with episodes of recurrent ischemic strokes often progressing to subcortical dementia.[206]

Recent work on the Drosophila RNA Helicase mahe, showed a crucial action of this gene in regulating Notch signalling in fruit-flies[207], thus suggesting a possible pathophysiological mechanism for explaining the radiological changes associated to DDX59 loss-of-function biallelic mutations. Importantly, our findings confirmed the association of DDX59 biallelic variants and digital/midline abnormalities with distinctive facial features and intellectual disability. However, the homozygous mutations reported by Shamseldin *et al*. (2013) were missense variants (c.1100T>G, p.Val367Gly; c.1600G>A, p.Gly534Arg) and the mutation recently described by Faily et al. (2017) affected a stop codon (c.1859G>T; p.*620Leuext*22) extending the protein product, whereas the homozygous mutation we identified in the present study lead to an early truncation of the protein, likely explaining the more severe phenotype including the complex heterogeneous neurological involvement of the patients.[197; 200] Interestingly, our functional analysis of the Drosophila model clearly showed that mahe null embryos exhibit severe neuronal phenotype in both peripheral as well as central nervous system, with defects in neurons during embryonic development consisting in highly disorganised neuron clusters and axonal projections. Severe loss or incomplete ventral nerve cord along with gaps in mahe mutant reflects loss of neurons and failure in proper neurogenesis. Also, hypomorphic mahe alleles also show shortened lifespan.

In conclusion, this study indicates a vital role of the RNA helicase mahe for nervous system development and function in Drosophila, and is also required to maintain normal lifespan. The same function must be conserved across species, thus supporting deleterious neurological consequences of DDX59 biallelic truncating mutations in humans. Further work remains to be done in fully understanding the role of this gene as well as other conserved DEAD-box RNA helicases in brain development and function.

4.4 De-novo and biallelic pathogenic variants in *NARS1* cause neurodevelopmental delay due to toxic gain of function and partial loss-of-function effect

4.4.1 Introduction

The attachment of tRNA to cognate amino acids is essential for protein translation. Aminoacyl-tRNA Synthetases (ARSs) are a group of enzymes encoded by ancient genes which are ubiquitously expressed and highly conserved.[208-210] These enzymes play a fundamental role in the esterification of proteinogenic amino acids to cognate tRNA. In total, thirty-seven genes encoding ARS enzymes have been described. Of these, twenty encode enzymes that function in the cytoplasm and the remainder relate exclusively to mitochondrial enzymes. Despite the essential canonical function and ubiquitous expression of ARS enzymes, mutations in these genes have been implicated in a variety of human diseases with both recessive and dominant inheritance patterns. [112; 209; 211; 212] These mutations result in neurological disorders, ranging from mild late-onset peripheral neuropathy to severe multi-systemic neurodevelopmental disorders (Supplementary Tables 3-5). [211-215]

Mutations in cytoplasmic ARS-encoding genes cause peripheral nervous system degeneration resulting in Charcot-Marie-Tooth neuropathies (*GARS1*, *AARS1* [MIM 601065]) and brain stem and spinal cord hypomyelination (*DARS1* [MIM 603084]). ARSs, and ARSs interacting genes, including *DARS1*, *RARS1* [MIM 107820], *AIMP1* [MIM 603605], and *AARS1*, have been implicated in neurodevelopmental disorders and epilepsies. Furthermore, mitochondrial *ARS2* mutations are often associated with leukoencephalopathy (*AARS2* [MIM 615889], *DARS2* [MIM 611105]) or pontocerebellar hypoplasia (*RARS2* [MIM 611524]). More recently, recessive mutations in *FARSA* [MIM 602918], *VARS1* [MIM 192150], *CARS1* [MIM 123859] and *TARS1* [MIM 187790], with subsequent partial loss of the ARS protein, have been linked to neurodevelopmental phenotypes. [216-220]

Modes of inheritance can be dominant or recessive; in cases such as *AARS1, YARS1* [MIM 603623]*, MARS1* [MIM 156560], *HARS1* [MIM 142810] and *GARS1,* both patterns can occur.[112]

The loss of function associated with mutations in ARSs is attributed to decreased aminoacylation efficiency or misfolding, causing protein instability with lower steady state levels[219]. However, in some cases (*GARS1, YARS1,* and *AARS1*) it has not been possible to ascribe the phenotype to a loss of primary aminoacylation.[221-223] Overall, the physiological functions of ARS genes and previously identified disease associations indicate an essential biological role for these proteins, implying that defects in all ARSs incur disease (Supplementary Table 2). [112]

Asparaginyl-tRNA^{Asn} is generated by *asparaginyl-tRNA synthetase* gene (*NARS1)* [MIM 108410, NM_004539.4] in a reaction involving two steps. NARS1 first catalyses the ATP-dependent activation of asparagine (Asn) into Asn~AMP with the release of pyrophosphate, and then transfers the activated Asn onto tRNA^{Asn} with the release of AMP (Figure 38). Here, we report the clinical phenotypes associated with *de-novo* dominant and biallelic, autosomal recessive mutations in *NARS1* in thirty-two affected individuals from twenty-one families. We provide genetic proof for these mutations and analyse their impact using individual cell lines, neural progenitor cells and molecular modelling.

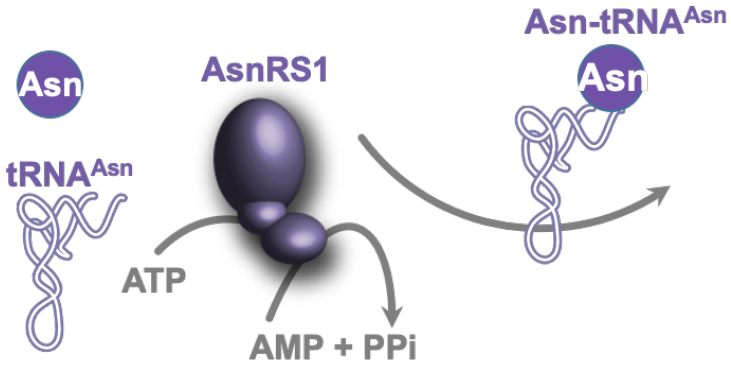


Figure 38. *AsnRS1* protein structure and function. Amino acid asparagine (Asn) is ligated to $tRNA^{Asn}$, catalysed by *AsnRS1* and ATP to produce Asn-tRNA (Asn), AMP and pyrophosphate (Manole *et al.*, 2020)**.**

In this study done together with Dr. Andreea Manole (postdoctoral researcher) and Dr. Emer O'Connor (neurologist), we described thirty-two individuals from twenty-one families, presenting with microcephaly, neurodevelopmental delay, seizures, peripheral neuropathy and ataxia, with *de-novo* heterozygous and biallelic mutations in the asparaginyl-tRNA-synthetase gene (*NARS1*). Through the work done in collaboration with Dr. Marisa Mendes (Amsterdam), Dr. Felix Distelmaier (Dusseldorf) and Dr. Rita Horvath (Cambridge) we demonstrated a reduction in NARS1 mRNA, protein expression and NARS1 enzyme levels in both individual fibroblasts and induced neural progenitor cells (iNPCs). In collaboration with Dr. Hubert Becker and Dr. Jean-Louis Mandel (Strasbourg), molecular modelling of the recessive c.1633C>T, p.Arg545Cys variant showed weaker spatial positioning and tRNA selectivity. We conclude that *de-novo* and biallelic mutations in *NARS1* are a significant cause of neurodevelopmental disease, where the mechanism for *de-novo* variants could be toxic gain of function and for recessive variants, partial loss-of-function.[5]

4.4.2 Results

Genetic findings

Using WES and homozygosity mapping as well as WGS done by our collaborators, we identified in total twenty-one families (F1-21) and thirty-two affected individuals (P1-32) with mutations in *NARS1* (Figure 39 shows *NARS1* variant schematic and Figure 41 illustrates pedigrees). Eight families had *de-novo* heterozygous variants; six had c.1600C>T, p.Arg534* (F1-6, P1-6); one had c.1525G>A, p.Gly509Ser (F7, P7); and one with c.965G>T, p.Arg322Leu (F8, P8). These variants were not present in our 652 normal brain series or in GnomAD databases.

Biallelic variants were found in thirteen families. Seven have homozygous c.1633C>T, Arg545Cys variants (F9-15, P9-23); one has homozygous c.50C>T, p.Thr17Met (F17, P26); and one has two siblings with homozygous c.32G>C, p.Arg11Pro (F16, P24 and 25). For compound variants, one family has two siblings with compound heterozygous c.1067A>C, p.Asp356Ala, c.203dupA, p.Met69Aspfs*4 (F19, P29 and P30). Two siblings had compound heterozygous c.1049T>C, p.Leu350Pro; c.1264G>A, p.Ala422Thr variants (F18, P27 and P28). There was one case with the compound heterozygous variant c.268C>T, p.Arg90*; c.394G>T, p.Gly132Cys (F20, P31), and a final individual with compound heterozygous c.1376C>T, p.Thr459Ile; c.178A>G, p.Lys60Glu (F21, P32). In GnomAD, c.1264G>A, p.Ala422Thr is present in 6 heterozygote individuals whilst c.1633C>T, p.Arg545Cys and c.50C>T, p.Thr17Met were present in 5 and 4 heterozygotes respectively. The c.100 A>T, p.Met34Leu, c.203dupA, p.Met69Aspfs*4, and c.1049T>C, p.Leu350Pro were absent, while c.32G>C, p.Arg11Pro was present in 1 individual. The c.1067A>C, p.Asp356Ala variant in family 19 was present in 264 heterozygotes, suggesting this variant may modify the phenotype and be pathogenic only when in trans with a severe variant such as c.203dupA, p.Met69Aspfs*4.

Data and Code Availability: The variants reported in this study have been submitted to the Leiden Open Variation Database and the accession numbers are: 668185, 668186, 668187, 668188, 668189, 668190, 668191, 668192, 668193, 668194, 668195, 668196, 668197 and 668198.

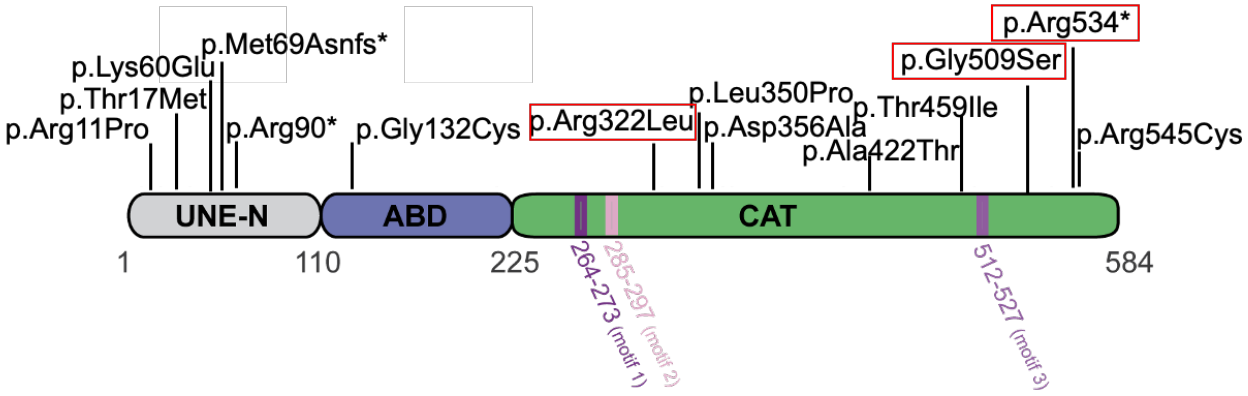


Figure 39. Schematic representation of human *ARS1* primary structure. Three main domains are depicted, the unique domain (UNE-N), the anticodon binding domain (ABD) and the catalytic domain (CAT). The nature and position of the mutants are shown above the primary structure, de-novo boxed in red, while the position of the domains are indicated below including motif 1 (involved in *AsnRS1* dimerization) and motifs 2&3 that form the active site (Manole *et al.*, 2020).

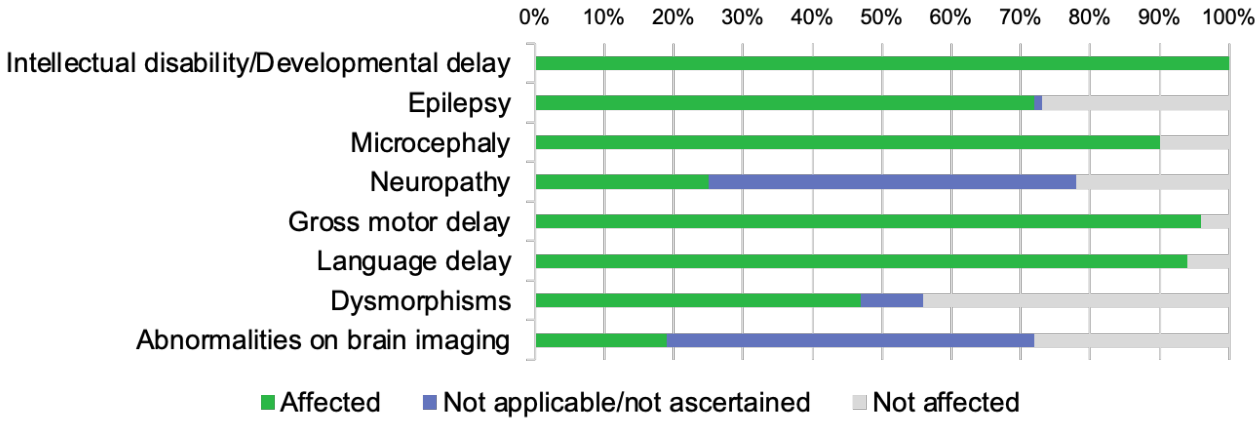


Figure 40. Bar graph summarizing proportions of various clinical findings affecting *NARS1* individuals (Manole *et al.*, 2020)**.**

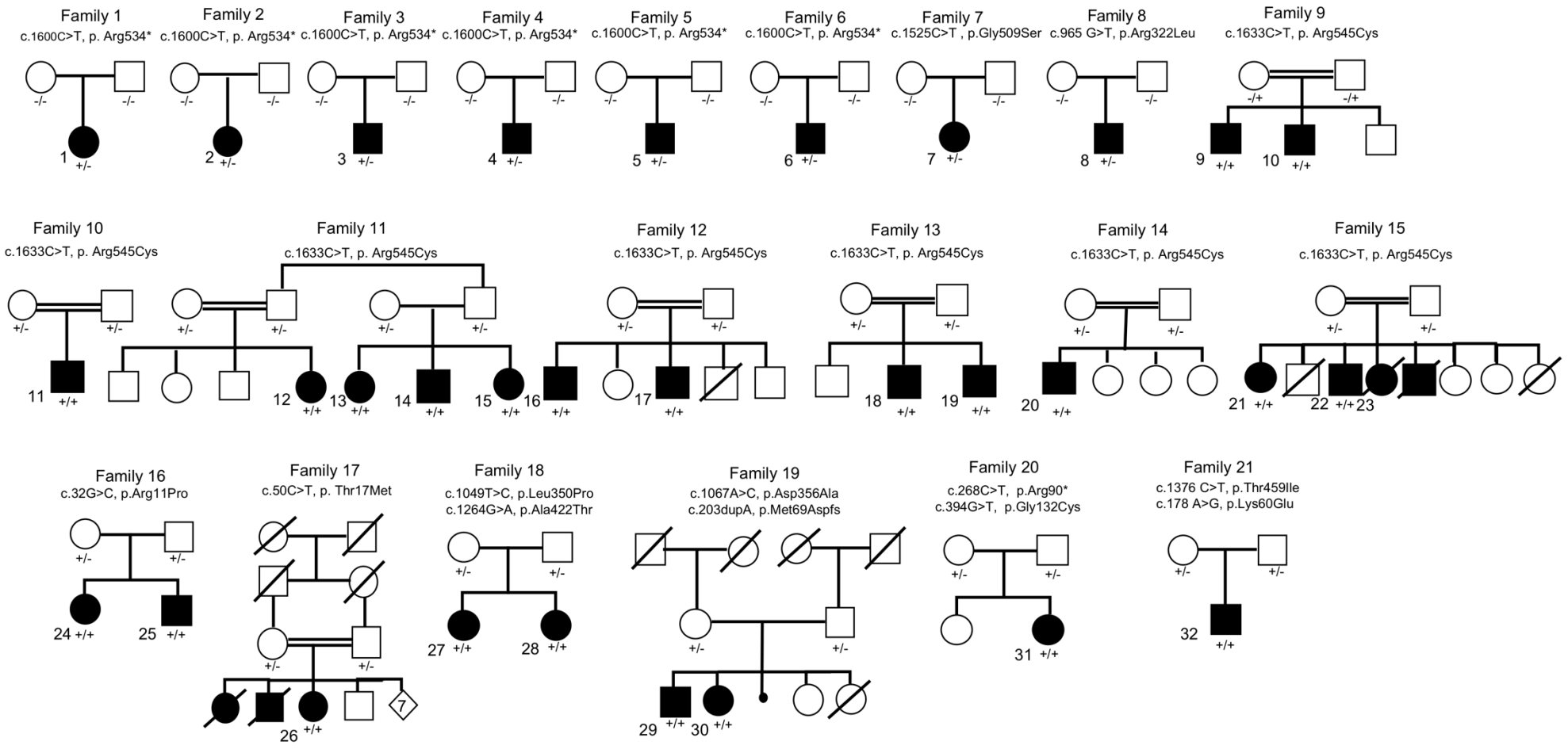


Figure 41. Pedigrees of the 21 families and 32 affected individuals identified in this study with *de-novo* and biallelic mutations in *NARS1*. Filled symbols represent affected individuals and double bars represent consanguinity in the family. -/-, +/-, +/+ represent wild-type, heterozygous and homozygous variants respectively (Manole *et al.*, 2020).

Clinical findings

Table 21 summarizes the core clinical features of affected individuals with *NARS1* defects and detailed tables and case reports are provided in the Appendix (Supplementary Tables 3-5). All individuals had global developmental delay (GDD) and intellectual disability, which varied from moderate to profound in severity. They had marked delay in language development. Motor development was also severely impaired with one individual never acquiring autonomous ambulation. Microcephaly was observed in the majority of cases (90%). These cases predominantly presented with primary microcephaly, however secondary microcephaly was also noted. Epilepsy was highly associated with the phenotype, affecting 23 cases (74.2%), with six individuals experiencing seizures below the age of one (Figure 40). The semiology of these attacks varied, with a mixture of partial, myoclonic and generalised tonic-clonic seizures described. An ataxic gait, poor balance and dysarthria were frequently detected on examination suggesting an additional neurodegenerative process, however, no structural abnormality of the cerebellum was observed on imaging. A demyelinating peripheral neuropathy occurred in eight individuals (25%) who had distal leg muscle atrophy. Dysmorphic features described included abnormal hands (*e.g.*, clinodactyly, foetal finger pad, 2-3 toe syndactyly, slender fingers) and/or feet (*e.g.*, small feet, toe syndactyly, slender feet). Upslanting palpebral fissures was the most common facial dysmorphism reported. A broad forehead, wide mouth, wide set teeth, and low set ears with overfolded helices were also described. Skeletal abnormalities including scoliosis, pronounced thoracic kyphosis and *pes-cavus* were also noted. Behavioural traits associated with the phenotype included impulsivity, stereotypies with repetitive speech / hand movements and selective feeding rituals.

Variant; Nucleotide, Protein	**c.1600C>T, p.Arg534***	**c.1525G>A, p.Gly509Ser**	**c.965G>T, p.Arg322Leu**	**c.1633C>T, p.Arg545Cys**	**c.32G>C, p.Arg11Pro**	**c.50C>T, p.Thr17Met**	**c.1049T>C c.1264G>A, p.Leu350Pro p.Ala422Thr**	**c.1067A>C c.203dupA, p.Asp356Ala p.Met69Aspfs*4**	**c.268 C>T c.394G>T, p.Arg90* p.Gly132Cys**	**c.1376 C>T, c.178 A>G, p.Thr459Ile, p.Lys60Glu**
Variant type	De-novo Heterozygous	De-novo Heterozygous	De-novo Heterozygous	Homozygous	Homozygous	Homozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous
Inheritance	AD de-novo	AD de-novo	AD de-novo	AR	AR	AR	AR	AR	AR	AR
Family	1-6	7	8	9-15	16	17	18	19	20	21
Affected Individual(s)	1-6	7	8	9-23	24-25	26	27-28	29-30	31	32
Ethnicity	European	UK	European	Pakistan / North India	Kosovo	Libya	German	Turkey	Canada	USA
Age at onset	Birth	Birth	Birth	Childhood	Childhood	Birth	Birth	Birth	Birth	Childhood
Consanguinity	No	No	No	Yes	No	Yes	No	No	No	No
Presentation	Severe GDD	Severe GDD	Severe GDD	Severe GDD	Seizures	Seizures	Mod GDD	Mod GDD	Severe GDD	Severe GDD
ID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Microcephaly	Yes	No	NA	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dysmorphic	Yes	Yes	Yes	Yes	No	NA	No	No	Yes	No
Seizures Affected Individuals	Yes 1,2,4,5,6	Yes	Yes	Yes 9,14,15,18,19, 21,22,23	Yes All individuals	Yes	Yes 27	Yes All individuals	Yes	Yes
Spasticity	Yes	No	Yes	No	Yes	NA	No	NA	No	Yes

Affected Individuals	3,4,6			Hypotonia in 9,10,16,17	24		Hypotonia		Hypotonia	
Neuropathy Affected Individuals	Yes 1, 2, 5	NA	NA	Yes 9,10,20	NA	NA	Yes	NA	NA	NA
Ataxia Affected Individuals	Yes All individuals	NA	Yes	Yes 9-12, 21	NA	NA	Yes	NA	Yes	Yes

Table 21. Summary of *NARS1* variants and clinical features of affected individuals. AD = autosomal dominant, AR = autosomal recessive, GDD = global development delay, ID = intellectual disability Mod = moderate, NA = not available.

Genotype-phenotype correlations

Family 16, with the homozygous variant p.Arg11Pro, had a particularly severe clinical picture comprising of severe developmental delay, progressive microcephaly, refractory seizures from infancy, and arrested myelination with pronounced cerebral atrophy on MRI. Otherwise, imaging was normal apart from microcephaly. There was no common structural change across all cases. Individuals with the de-novo p.Arg534* variant showed severe microcephaly. In one family with this variant (F6), mild atrophy was observed (Supplementary Table 3-5 and Figure 42).

Individuals homozygous for p.Arg545Cys demonstrated hypotonia and predominantly distal weakness. Spasticity was observed in individuals with the p.Arg11Pro or *de-novo* variants.

A demyelinating polyneuropathy was documented in cases homozygous for the p.Arg545Cys variant (P9,10,20), and, in one case, this was confirmed with a sural nerve biopsy (F9, P9). It was also described in cases with the *de-novo* c.1600C>T, p.Arg534* variant (P1,2,5) and the family with the compound heterozygous c.1049T>C, p.Leu350Pro, c.1264G>A, p.Ala422Thr variant (F18, P27,28).

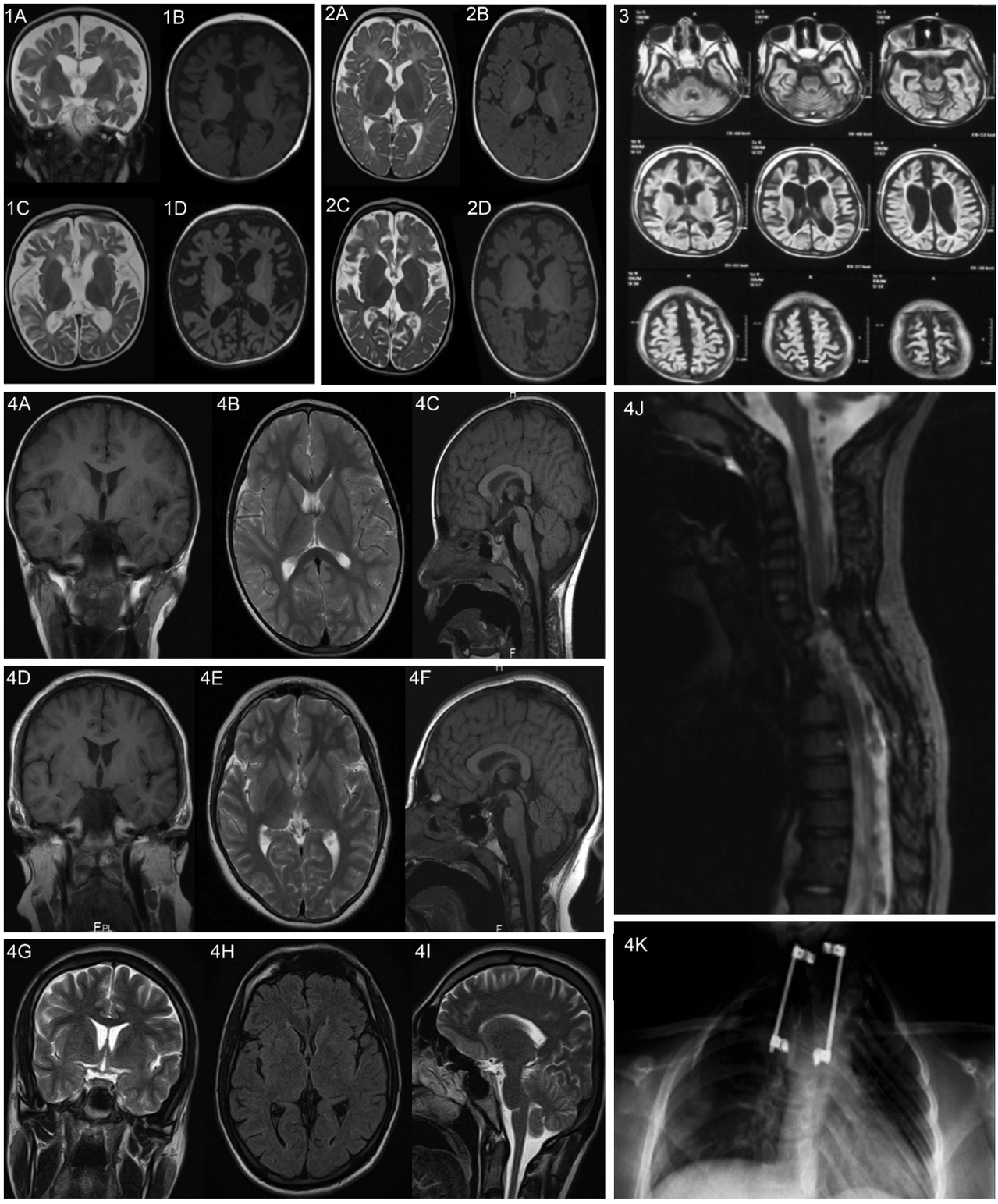


Figure 42. Radiological findings of individuals in our cohort. Set 1: Case of homozygous c.32G>C, p.Arg11Pro variant. Upper row images [coronal T2-WI (1A) and axial T1-WI (1B)] at the age of 10 months show severely delayed myelination and fronto-temporal atrophy. Lower row images [axial T2-WI (1C) and axial T1-WI (1D)] repeated at the age of 18 months shows progressive and global brain atrophy

with an emerging pattern of severe hypomyelination. **Set 2:** A further case of homozygous c.32G>C, p.Arg11Pro variant. Upper row images [axial T2-WI (2A) and axial T1-WI (2B)] at the age of 8 months show mild fronto-temporal underdevelopment and severely delayed myelination. Lower row images [axial T2-WI (2C) and axial T1-W1 (2D)] repeated at the age of 2 years shows progressive and global brain atrophy along with severe hypomyelination. **Set 3:** Case of homozygous p.Thr17Met variant. Axial FLAIR images at the age of 9 months shows global atrophy involving the cerebral and cerebellar hemispheres along with severe hypomyelination. **Set 4:** MRI images of an individual with the homozygous c.1633C>T, p.Arg545Cyc variant. Coronal T1-WI (4A), axial T2-WI (4B) and sagittal T1-WI (4C) at the age of 4 years, coronal T1-WI (4D), axial T2-WI (4E) and sagittal T1-WI (4F) at the age of 11 years as well as coronal T2-WI (4G), axial FLAIR (4H) and sagittal T2-WI (4I) at the age of 20 years demonstrating normal intracranial appearances across the 3 different ages. This individual had an upper thoracic scoliosis which was operatively corrected at the age of 4 demonstrated on the sagittal T2-WI of the spine (4J) and frontal projection radiograph of the chest/thoracic spine (4K) (Manole *et al.*, 2020).

Pathogenicity of *NARS1* variants

NARS1 is intolerant to loss of function (missense variants constraint is Z=0.87). We identified *de-novo NARS1* mutations in eight families (F1-8, P1-8) with similar phenotypes. A variant at codon 534 recurred in six families (F1-6, P1-6). The two other de-novo variants were located at codon 322 and 509. The c.1600C>T, p.Arg534* variant is located 15 amino acids (AA) from the end of the 548 AAs protein, representing a potential hotspot for pathogenic mutations. Arginine at codons 534 and 545 are universally conserved in AsnRS1 from all three major taxonomic groupings, implying a significant structural or functional role.

The homozygous p.Arg545Cys variant was evident in seven recessive families affecting the same C-terminal catalytic stretch as c.1600C>T, p.Arg534*. Given its proximity to residue 534, c.1633C>T, p.Arg545Cys may have a comparable mechanistic effect to c.1600C>T, p.Arg534*.

The c.1067A>C, p.Asp356Ala variant was found in *trans* with the only recessive truncating allele observed thus far at c.203dupA; p.Met69Aspfs*4 (P29 and P30). Two missense variants (c.965G>T, p.Arg322Leu and c.653T>C, p.Asn218Ser) were found in P8, however, as c.965G>T, p.Arg322Leu occurred *de-novo*, it could not be determined whether these variants were in *cis* or *trans*. Moreover, the frequency of c.653T>C, p.Asn218Ser in the GnomAD database (78 heterozygotes) suggests it is unlikely to be associated with a severe phenotype, leaving c.965G>T, p.Arg322Leu as the most likely disease-causing variant. The p.Arg322 residue is essential for enzymatic activity and therefore is predicted to cause impaired enzyme activity (Pr. H. Becker, personal communication). Both the c.50C>T, p.Thr17Met and c.32G>C, p.Arg11Pro variants are in the N-terminal UNE-N appended domain of AsnRS1, specific to eukaryotes, recently shown to have chemokine activity.[224]

Functional characterisation

Western blotting and RT-PCR

Given the potential loss of function in homozygous *NARS1* individuals, we investigated gene expression levels by semi-quantitative PCR and protein levels by Western blotting of *AsnRS1* from lymphoblasts and fibroblasts from families with p.Arg545Cys, p.Thr17Met, p.Asp356Ala; p.Met69Aspfs*4, p.Arg11Pro, and show that *AsnRS1* cDNA expression and protein levels are reduced (Figure 4).

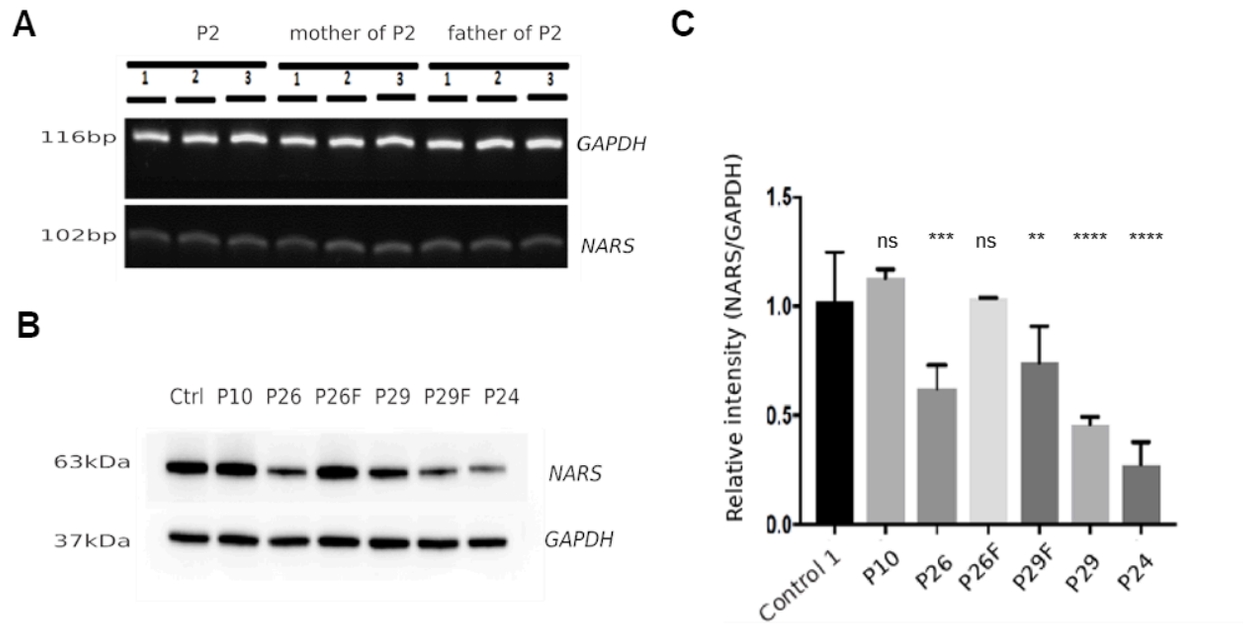


Figure 43. Protein levels of *AsnRS1* are clearly reduced in individual-derived cells. **A.** RT-PCR of the *de-novo* c.1600C>T, p.Arg534* variant in P2 and parents **B**. Western blotting and **C.** quantification graph of the Western blot of *NARS1* individuals with mutations compared with controls. Ctrl = control, P10 = homozygous c.1633C>T, p.Arg545Cys, P26 = homozygous c.50C>T, p.Thr17Met, P29 = compound heterozygous c.1067A>C, p.Asp356Ala and c.203dupA, p.Met69Aspfs*4 (F denotes father of individuals), P24 = homozygous c.32G>C, p.Arg11Pro (Manole *et al.*, 2020).

Induced Neural progenitor cells

iNPC colonies were produced and isolated from fibroblasts from P26 (c.50C>T) and P29/30 (c.1067A>C, c.203dupA). From the isolated colonies gene expression was determined by qPCR to select iNPC populations which presented decreased expression of fibroblast markers *COL1A1*, *COL3A1*, *TWIST2* and *DKK3*, as well as increased NPC markers *NES*, *SOX1* and *MSI1*, and the iNPC population was expanded to be subsequently used for RNA sequencing (RNAseq). RNAseq showed normal *NARS1* expression in iNPCs from affected individuals carrying the c.50C>T and c.1067A>C

mutations, and decreased expression of the c.203dupA *NARS1* allele in the P29 cells. Interestingly, iNPCs from affected individuals show increased expression of several other ARSs (*DARS1, GARS1, RARS1, SARS1, TARS1, WARS1, YARS1*) (Figure 44). This could be explained by the fact that the *NARS1* mutant(s) are inducing the Integrated Stress Response (ISR), which activated a number of ARS genes as a result of the loss of function homozygous recessive variants. Impaired synthetase function may reduce the amount of charged tRNA available for translation elongation, with a possible increase in the levels of uncharged tRNA. Uncharged tRNAs produced as a result of amino acid deprivation have been reported to bind GCN2, leading to the activation of the ISR.[225] Analysis of the cellular pathways (Reactome, GO) associated with genes with significantly altered mRNA levels showed that upregulated genes were enriched (adjusted p-value <0.01) for pathways heavily associated with protein translation and processing such as ER and Golgi protein processing and ribosomal homeostasis. In addition, increased action of VEGFR1/2 (upregulated by ATF4, which is one of the key transcription factors in the ISR) was suggested.

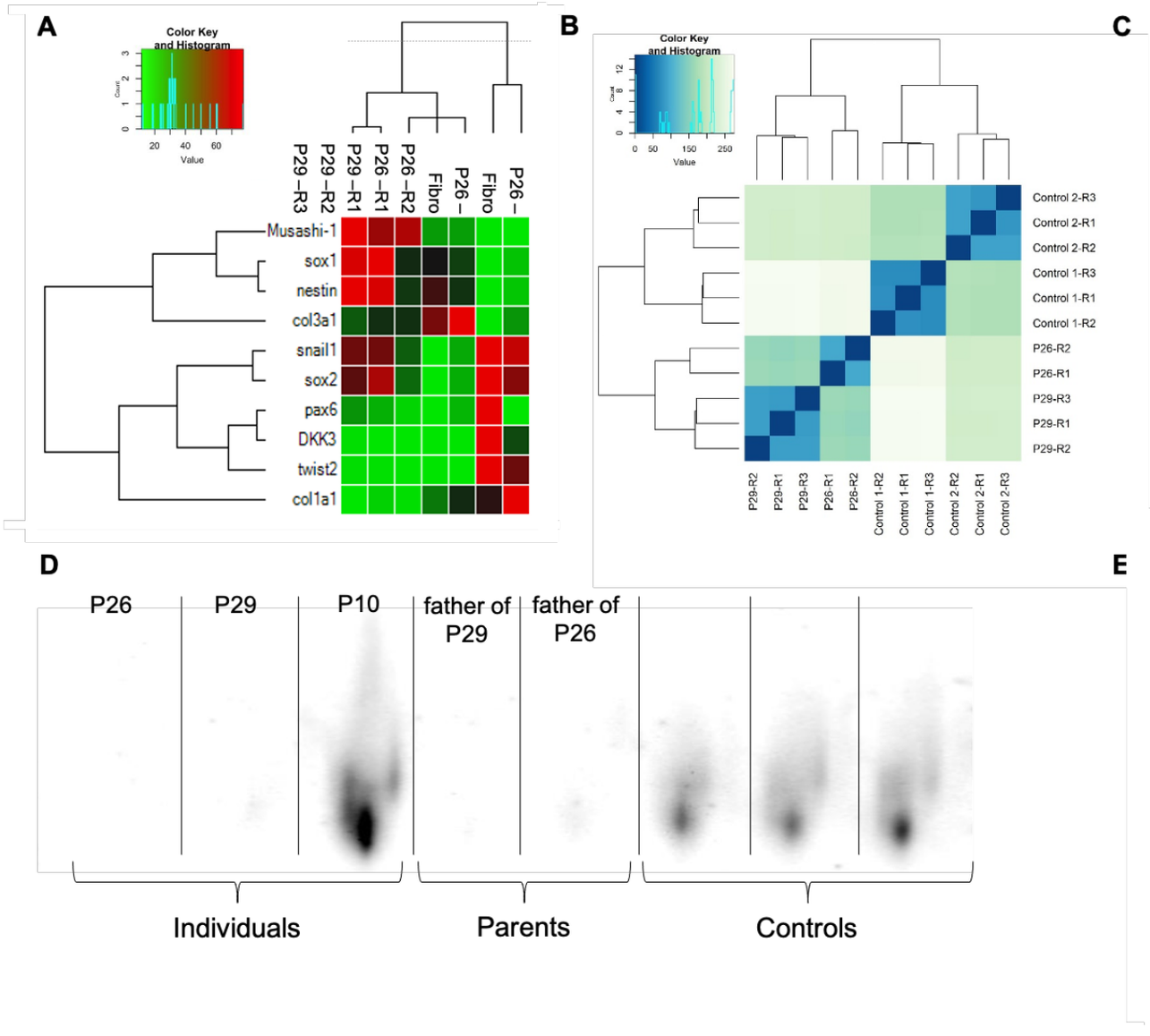

A
Color Key and Histogram
Count
Value
P29 –R3
P29 –R2
P29 –R1
P26 –R1
P26 –R2
Fibro
P26 –
Fibro
P26 –
Musashi-1
sox1
nestin
col3a1
snail1
sox2
pax6
DKK3
twist2
col1a1
B
Color Key and Histogram
Count
Value
Control 2-R3
Control 2-R1
Control 2-R2
Control 1-R3
Control 1-R1
Control 1-R2
P26-R2
P26-R1
P29-R3
P29-R1
P29-R2
P29-R2
P29-R1
P29-R3
P26-R1
P26-R2
Control 1-R2
Control 1-R1
Control 1-R3
Control 2-R2
Control 2-R1
Control 2-R3
C
-10*log10 padj
D
P26
P29
P10
father of P29
father of P26
Individuals
Parents
Controls
E
Relative expression

Figure 44. BNPAGE and iNPC RNA-sequencing. A. iNPC populations from P26 (c.50C>T, p.Thr17Met) and P29 (c.203dupA, p.Met69Aspfs*4; c.1067A>C, p.Asp356Ala) exhibit increased expression (red coloured) of most iNPC markers (sox1, sox2, nestin, snail1, pax6, DKK3, twist2, Musashi-1) compared to fibroblast markers (green coded) as measured by qPCR, shown with hierarchal clustering. R1-R3 represent replicates. **B.** Heatmap with hierarchal clustering generated using all gene counts from RNAseq distinction of control (Ctrl1 a-c, Ctrl2 a-c) and individual-derived (P26 a-b, P29 a-c) iNPCs. **C.** Volcano plot showing log2 of fold change in NARS mutant iNPCs compared to controls and - log10 (adjusted p value). **D.** BN-PAGE-Western blot showing reduced levels of the AsnRS1 dimer in individuals P26, P29 and fathers compared to control, but not for individual P10. **E.** Quantification of BN-PAGE-Western blot AsnRS1 dimer formation, showing significantly (*** $p<0.001$) reduced levels of the AsnRS1 in individuals P26, P29 and fathers compared to control but not change for P10. P26=homozygous c.50C>T, p.Thr17Met, P29=c.203dupA, p.Met69Aspfs*4; c.1067A>C, p.Asp356Ala, P10= c.1633C>T, p.R545C, Father of P26=heterozygous c.50C>T, p.Thr17Met, Father of P29= c.1067A>C, p.Asp356Ala (Manole *et al.*, 2020, data generated by Matthew Jennings)**.**

Blue-native polyacrylamide gel electrophoresis (BN-PAGE)

Examining the amount of the AsnRS1 dimer complex, we showed that P26 (carrying the p.Thr17Met variant) and P29 (carrying the p.Asn356Ala and p.Met69Aspfs*4 variants) had severely reduced dimer formation compared to healthy controls (Figure 44E). The individual's unaffected parents carrying one heterozygous mutation each also appeared to show decreased level of the AsnRS1 dimer. P10 (p.Arg545Cys) showed comparable AsnRS1 dimer amount to healthy controls. Similar to most other disease-associated ARSs, AsnRS1 functions as a class II homodimer[226]. The decreased AsnRS1 dimer

formation observed in fibroblasts from P26 and P29 shown by BN-PAGE accounts for the apparent deficit in aminoacylation capacity, despite showing no consistent decrease in AsnRS1 monomer expression. This is further supported by the molecular model simulation (Figure 45) that predicts an unstable dimer for the p.Asn356Ala mutant since this substitution is located at the interface between the two AsnRS1 monomers. Further predicted effects of *NARS1* missense mutations are shown in Supplementary Table 7.

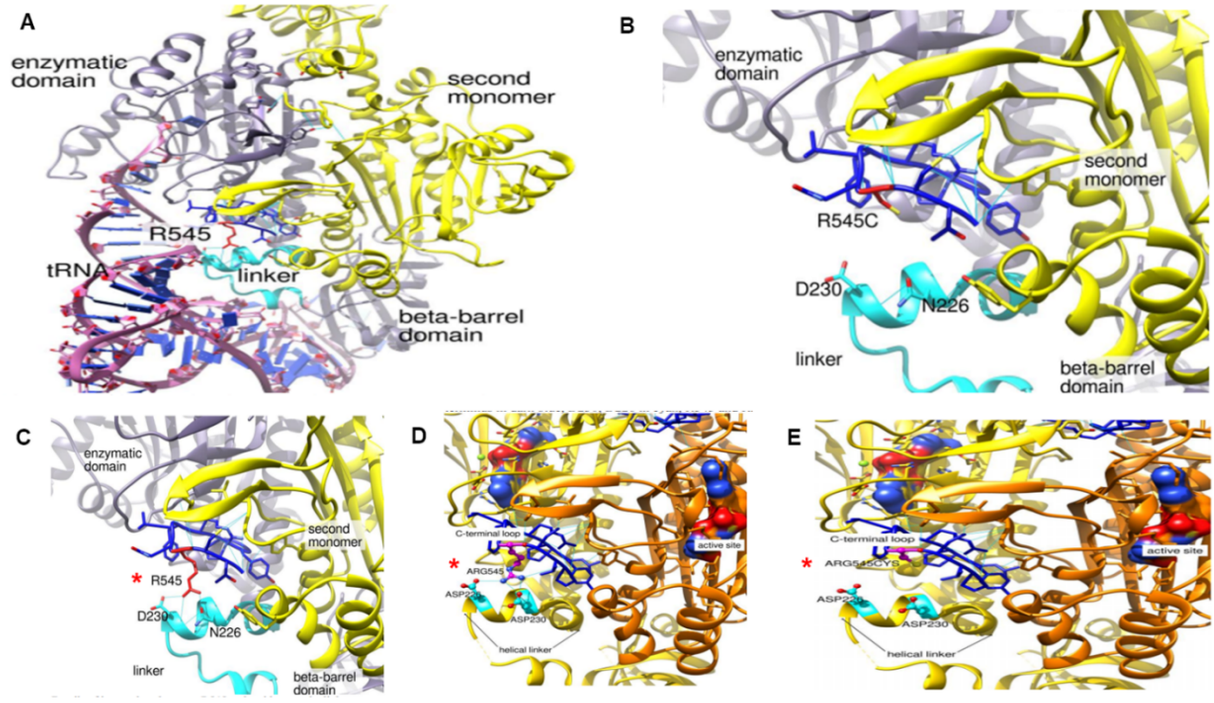


Figure 45. Molecular modelling of the *NARS1* Arg545Cys homozygous variant. The crystal structure is based on *B.malayi AsnRS1*. *AsnRS1* is a homodimer; one *AsnRS1* monomer given in yellow and one in orange. Analog of the transition state presented in the surface representation, C-terminus in dark blue, Asp230, Asp226 in cyan, Arg545 and Arg545Cys in magenta. **A.** Interaction between *AsnRS1* and tRNA with residues on the helical linker. **B-E**. Zoom in on the C-terminus helical linker region, **B and E.** show loss of molecular interaction and folding of the *AsnRS1* protein with the p.Arg545Cys variant (*) (Manole *et al.*, 2020, data generated by Hubert Becker and Jean-Louis Mandel).

ARS enzymatic assays

In comparison with controls, AsnRS1 enzymatic activity in individual-derived fibroblasts and lymphoblasts with *NARS1* variants showed a reduction. The most dramatic decrease was observed for P2 (carrying the *de-novo* p.Arg534* variant) and the mildest decrease was observed for P24 (homozygous for the p.Arg11Pro) [80% of the controls]. Amino acid residue Arg11 is located in the 5' end of the non-canonical UNE-N domain (Figures 46 and 47), recently shown to elicit cell migration of human immune cells via migration of CC chemokine receptor 3 (*CCR3*) in an autoimmune disease associated with ARS genes.[224]

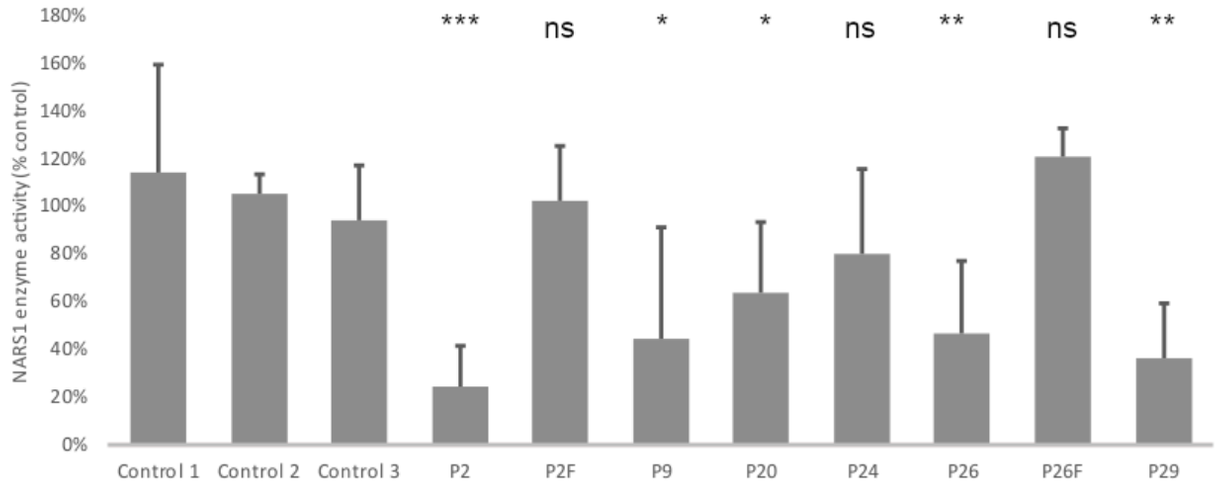


Figure 46. Reduced asparaginyl-tRNA synthetase activity in individuals cell lines carrying homozygous variants c.1600C>T, p.Arg534* (P2), c.1633C>T, p.Arg545Cys (P9, P20), c.32G>C, p.Arg11Pro (P24), c.50C>T, p.Thr17Met (P26) and c.1067A>C, p.Asp356Ala / c.203dupA, p.M69Aspfs*4 (P29) *NARS1* variants in comparison to the average of 3 unrelated fibroblast cell lines (all cell lines are fibroblast except P2 which is a lymphoblast cell line, control values for lymphoblast are similar to fibroblasts); n=9, p value FDR<0.01 (Manole *et al.*, 2020, data generated by Marisa Mendes)**.**

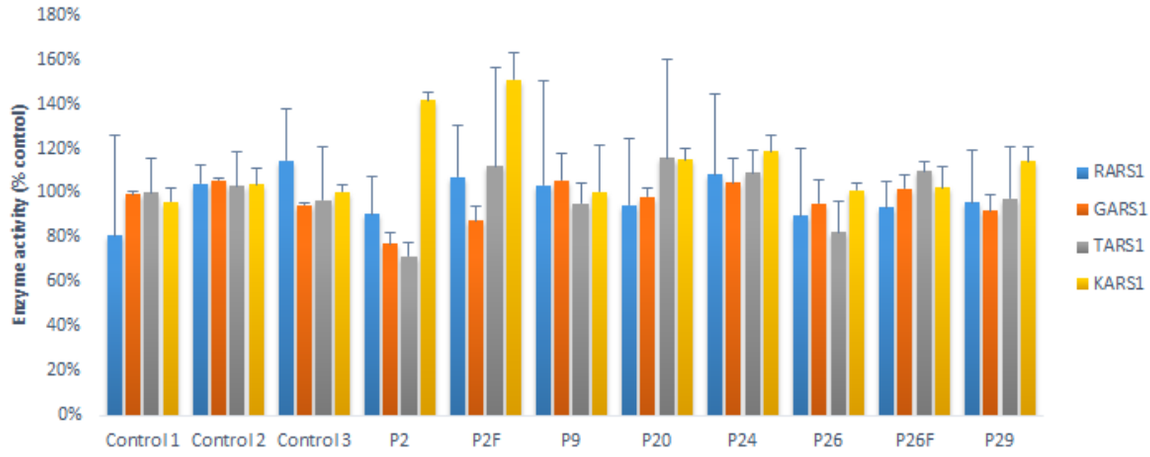


Figure 47. Individual derived fibroblast and lymphoblast cells displaying significant activity of threonyl-tRNA sythetase; lysyl-tRNA synthetase; glycyl-tRNA synthetase and arginyl-tRNA synthetase in individuals' and parents' cells (c.1600C>T, p.Arg534* (P2), c.1633C>T, p.Arg545Cys (P9, P20), c.32G>C, p.Arg11Pro (P24), c.50C>T, p.Thr17Met (P26) and c.1067A>C, p.Asp356Ala / c.203dupA, p.M69Aspfs*4 (P29); TARS: Threonyl-tRNA sythetase; KARS: lysyl-tRNA synthetase; GARS: Glycyl-tRNA synthetase and RARS: arginyl-tRNA synthetase were measured as control for AsnRS1 activity (Manole *et al.*, 2020, data generated by Marisa Mendes)**.**

4.4.3 Discussion

We identified *de-novo* heterozygous and biallelic mutations in the cytoplasmic *NARS1* in thirty-two individuals with a neurodevelopmental phenotype. Mutations included de-novo and recessive mutation hotspots at amino acid residue p.Arg534* and p.Arg545 respectively, in the last 40 amino acids of the protein. Two homozygous variants identified at the 5' end, p.Arg11 and p.Thr17, were associated with a severe clinical phenotype. Other mutations in *NARS1* were spread throughout and did not cluster in any particular region of the gene.

The clinical phenotypes associated with homozygous variants c.32G>C, p.Arg11Pro and c.50C>T, p.Thr17Met correlate with reduced expression and may reflect impaired protein stability as suggested by the structural modelling of c.1633C>T, p.Arg545Cys (Figure 45). Interestingly, MRI imaging of individuals carrying the c.32G>C, p.Arg11Pro and c.50C>T, p.Thr17Met variants showed atrophy and white matter abnormalities. In contrast, no such changes were identified in individuals with the Arg545Cys variant. The clustering of mutations and associated phenotypes at the N- and C-terminus suggest these regions are functionally important and disrupt the protein homodimer and ATP-binding/catalytic domain in *NARS1*. These two variants produced elevated AsnRS1 enzyme activity, which can be attributed to their location in the N-terminal extension domain. This domain has additional non-translational functions which enables enzymatic activity of the modified protein. By examining protein expression, protein synthesis, and the aminoacetylation activity, it is clear that the non-translational functions of such ARS proteins, regulated by the newly-evolved appended domains such as UNE-N, don't seem necessary for ARS activity (Figures 46, 47).

Our functional data, including fibroblasts and iNPCs transcriptomics, suggest that the majority of *NARS1* mutations cause a loss of the enzymatic protein by reduced expression and disruption of the homodimer formation. This results in abnormal protein synthesis and processing with a compensatory increase in expression of other ARSs (Figures 43, 44, 46, 47). The increased action of VEGFR1/2 was of interest considering the reported actions of other ARSs, as in *GARS* and *EPRS* via GAIT complex, *SARS1*, *TARS1* and mini *WARS1* on VEGF-related signalling.[227] Pathways associated with downregulated genes were typically associated with cell cycle progression, DNA repair and replication such as G2/M checkpoint, homology directed repair and telomere maintenance pathways, suggesting that this alteration of cellular proliferation could be a result of decreased protein synthesis (Figures 48-52). In general, the mutations in

NARS1 resulted in loss of function in both studied iNPC cell lines (P26 and P29) leading to a transcriptomic signature of induced ISR, upregulation of protein translation and processing in the ER and Golgi, and altered ribosomal homeostasis. This is similar to other studies in cells with reduced aminoacylation activity in disease-associated mutations in other cytosolic ARSs.[212; 227; 228]

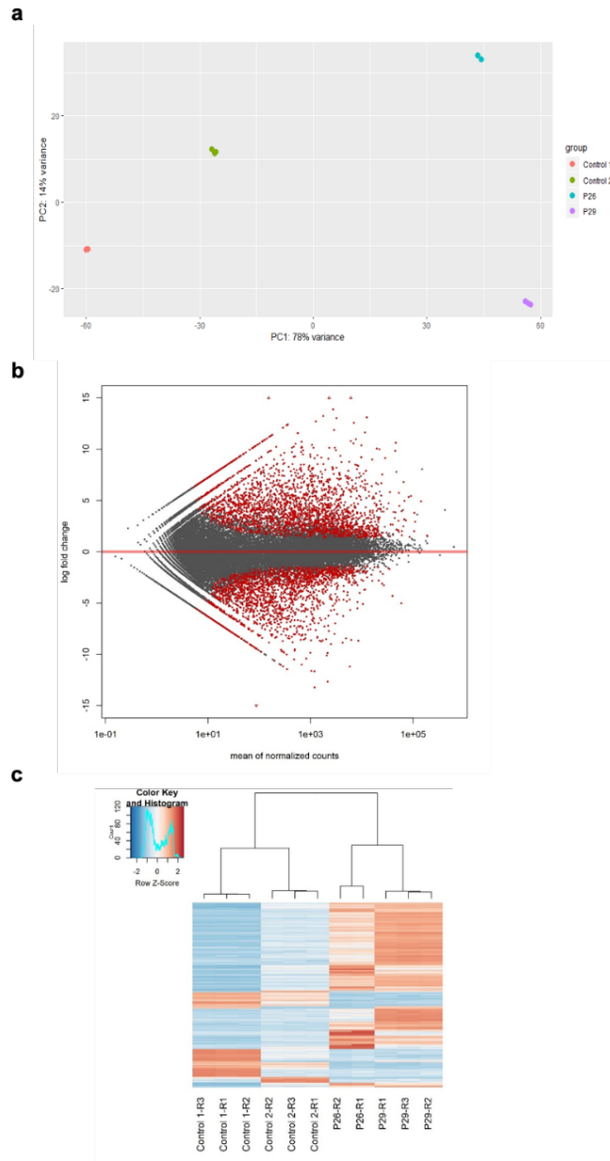


Figure 48. Transcriptomic gene expression analysis A. Principle component analysis of transcriptomic gene expression levels. B. MA plot showing normalised count number versus fold change, significantly differentially expressed gene shown in red. C. Cluster heatmap of top 750 most differentially expressed normalised gene counts (Manole *et al.*, 2020, data generated by Matthew Jennings).

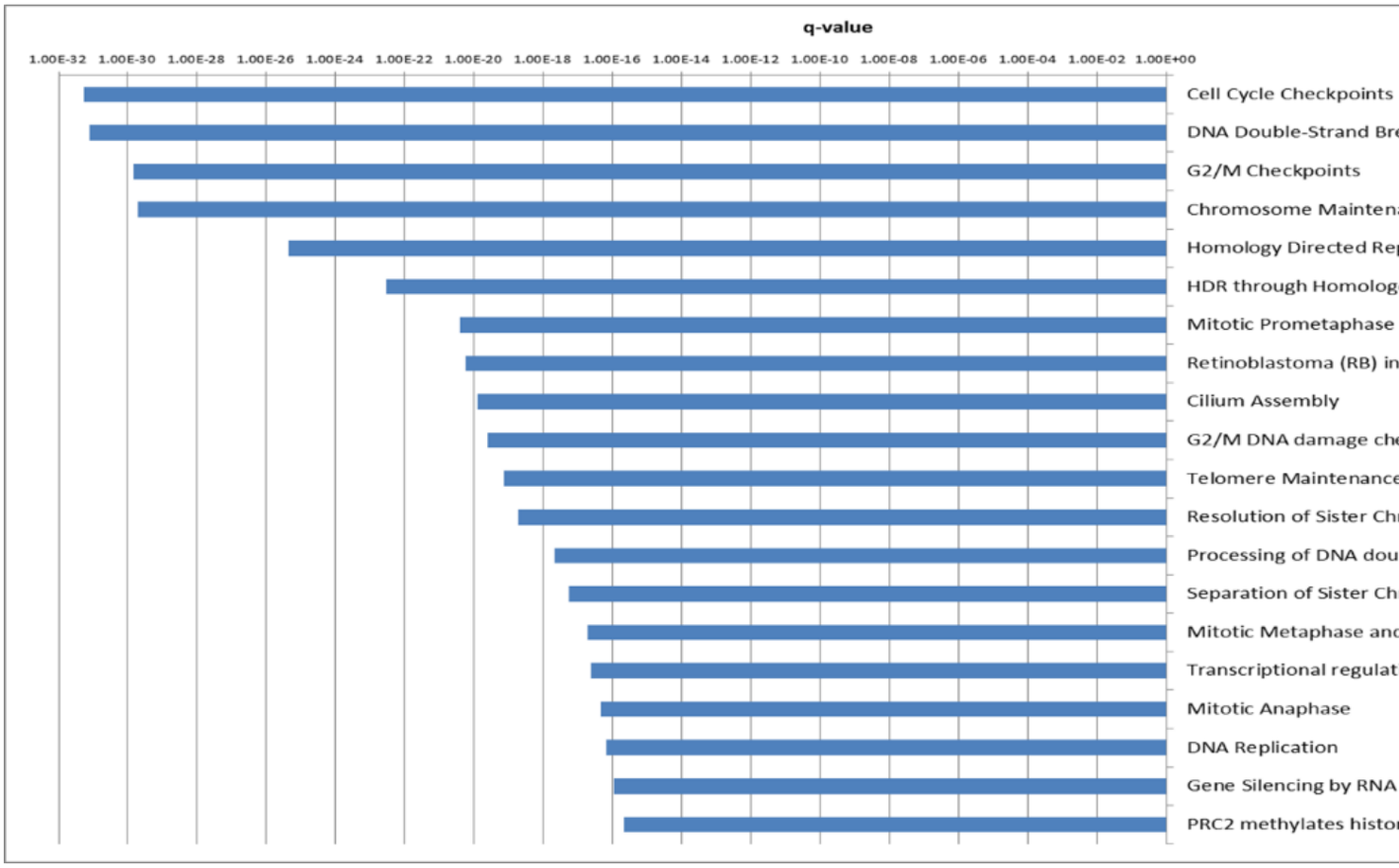


Figure 49. Supplementary iNPC data of top 20 (by significance) pathways associated with downregulated

et al., 2020)**.**

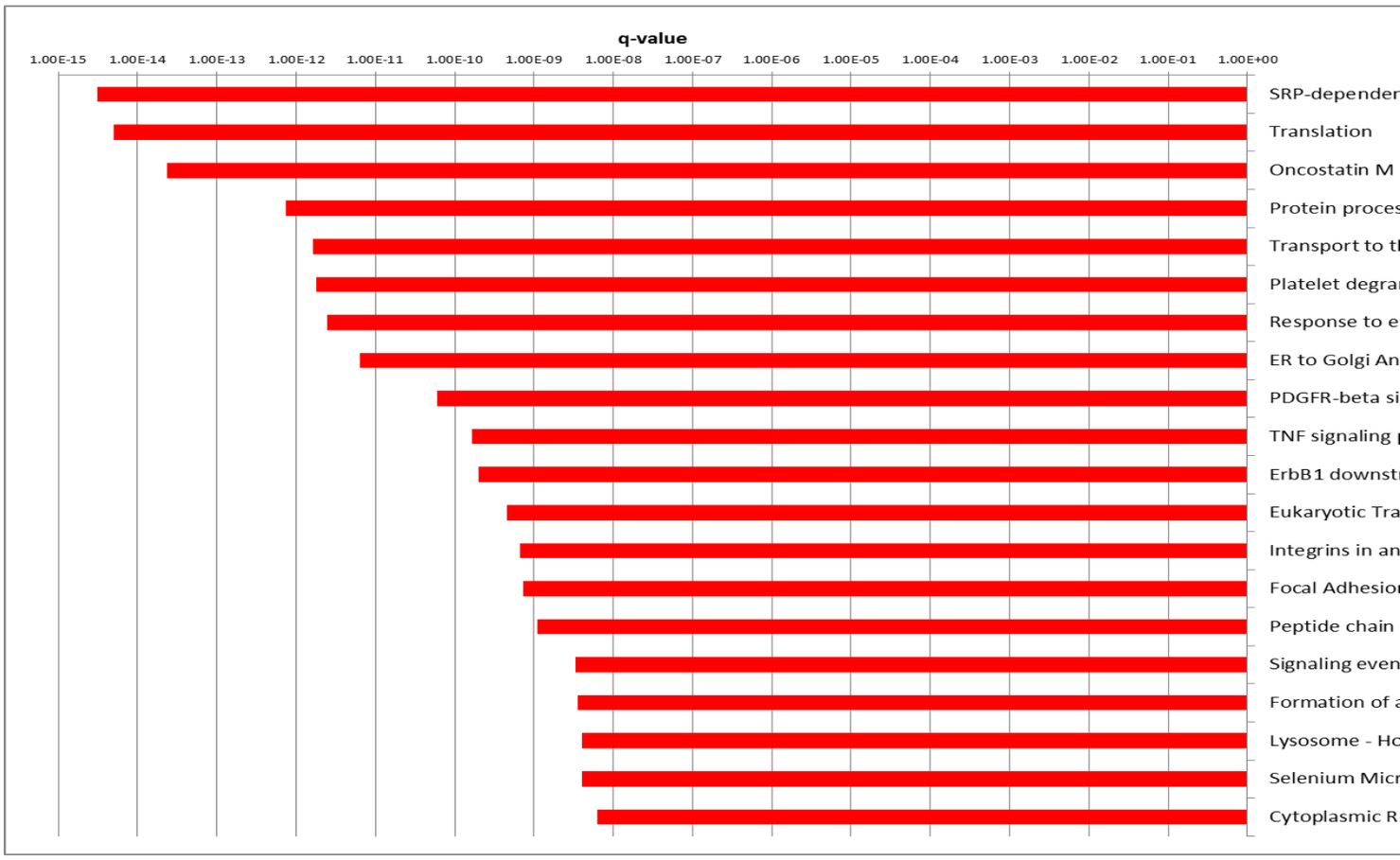


Figure 50. Supplementary iNPC data of top 20 (by significance) pathways associated with upregulated ge

al., 2020)**.**

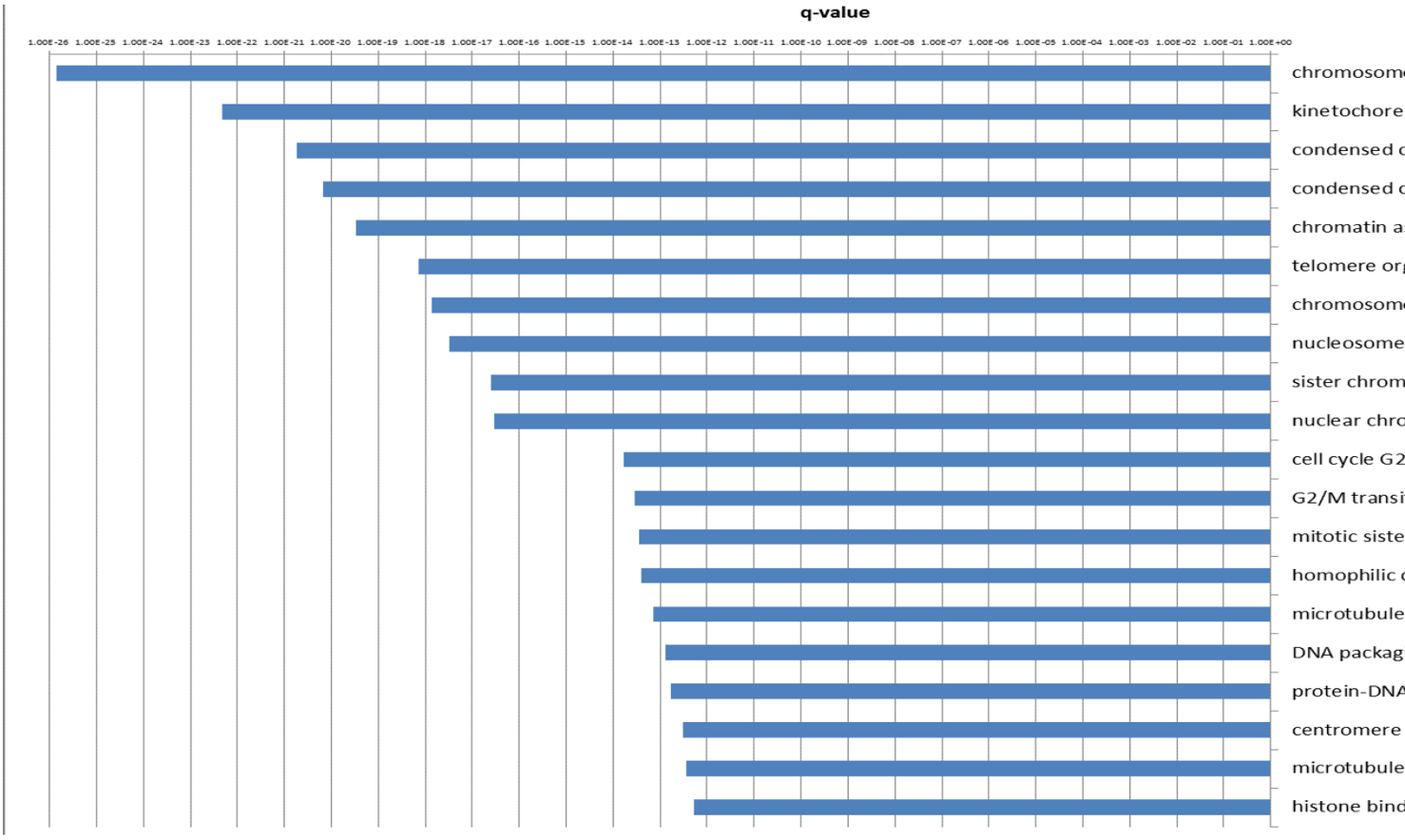


Figure 51. Supplementary iNPC data of top 20 (by significance) GO terms associated with downregulated

et al., 2020)**.**

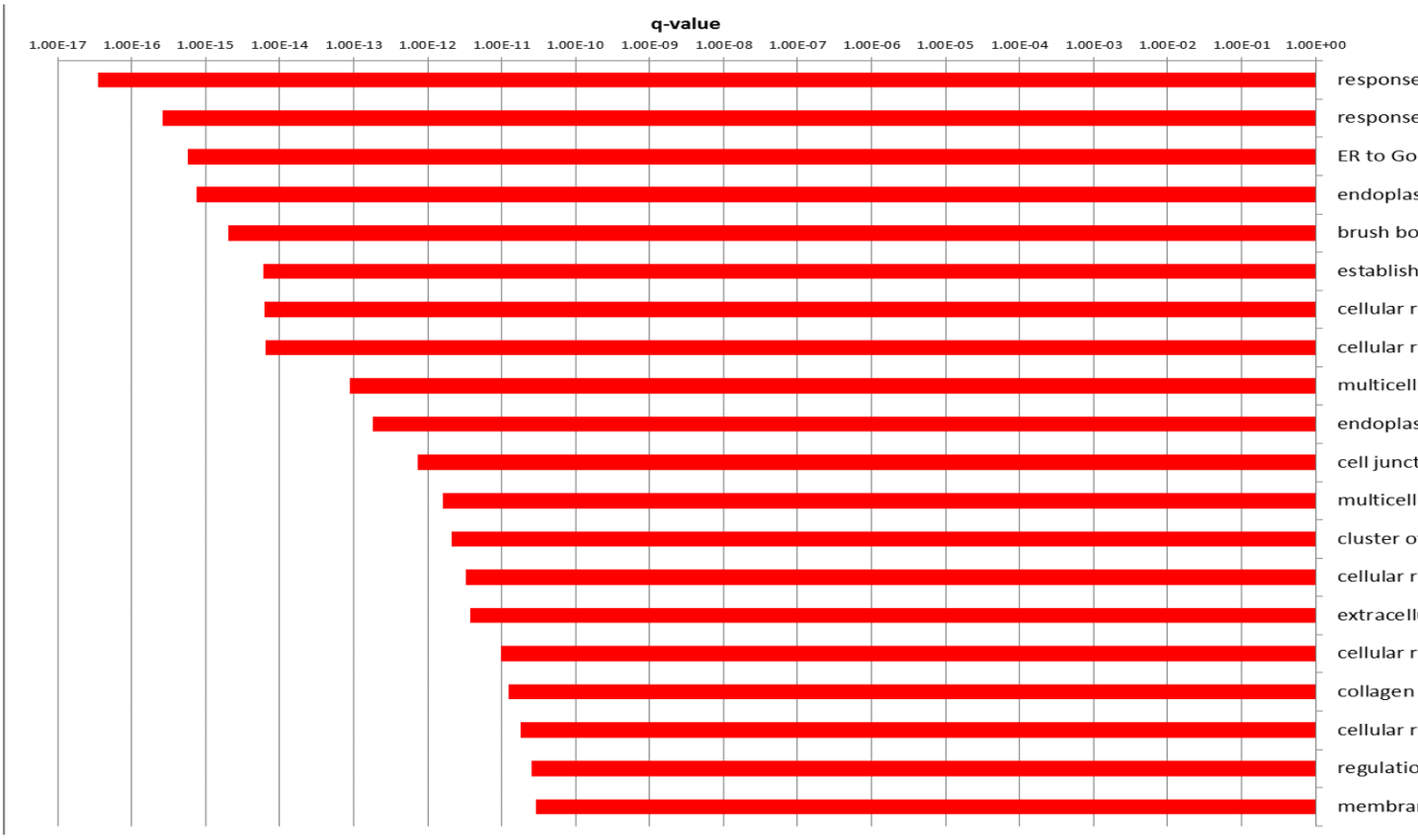


Figure 52. Supplementary iNPC data of top 20 (by significance) GO terms associated with upregulated ge

In summary, downregulated genes have been associated with DNA replication, whereas upregulated gene synthesis and processing. NARS1 gene expression was unaltered (p=0.45), while other tRNA synthetase WARS, TARS, YARS, GARS, SARS). LARS was downregulated & others were not altered. In the compou frameshift mRNA (p.Met69Asp*fs*) is only expressed at a 1:9 ratio compared to the allele with a single base also confirmed by Western blotting showing decreased AsnRS1 expression (Manole *et al.*, 2020)**.**

The common homozygous c.1633C>T, p.Arg545Cys variant in the Western blot (Figure 43) and yeast model done in collaboration with the Prof. Jurg Bahler lab (UCL) showed near normal protein levels and an increased yeast growth suggestive of a gain of function mechanism (Figure 53). However, protein modelling of this variant demonstrated loss of the helix linker indicating reduced tRNA interaction and catalytic activity. This loss of function effect was evidenced by the reduced aminoacylation activity to 40% compared to controls (Figures 46, 47). This effect could potentially be more harmful for cells of the nervous system than a unicellular organism. One of the possibilities is that the *NARS1* mutant mischarges a tRNA in the human cells that might be less conserved in fungi. Thus, the mischarging would be reduced in yeast, hence the better growth without side effects. Molecular modelling has shown that the p.Arg545Cys variant lies within a region that probably interacts with the sugar-phosphate backbone of the tRNA (at positions 68-69), close to the active site of the enzyme.[229] Replacing the bulky arginine with a cysteine does not seems to perturb the enzyme's overall structure (Figure 45). However, by disrupting the tRNA-enzyme contact, this variant may alter the enzyme selectivity towards tRNA, decreasing the overall affinity for tRNA. *AsnRS1* enzyme activity for individuals homozygous for this variant (P9 and P20) showed decreased activity (Figures 46, 47). Similarly, the *de-novo* c.1600C>T, p.Arg534* variant, located adjacent to the end of the protein terminus, has been shown to have a gain-of-function effect that interferes with normal protein function. It is likely a protein lacking the 15 amino acids containing the ATP-binding domain is produced. This region is important for the enzymatic function and escapes mRNA decay as shown by semi-quantitative RT-PCR from family 2 (Figure 43A). We further developed a zebrafish model of this variant (Figure 54) in collaboration with Dr. Dagan Jenkins (UCL) where a dominant negative effect on the wild-type allele has been shown, causing a dose-dependent phenotype, specifically cyclopia and

gastrulation defects at 200-500pg. Similar cyclopia defects in zebrafish were reported for microcephaly gene *ORC1.*[230]

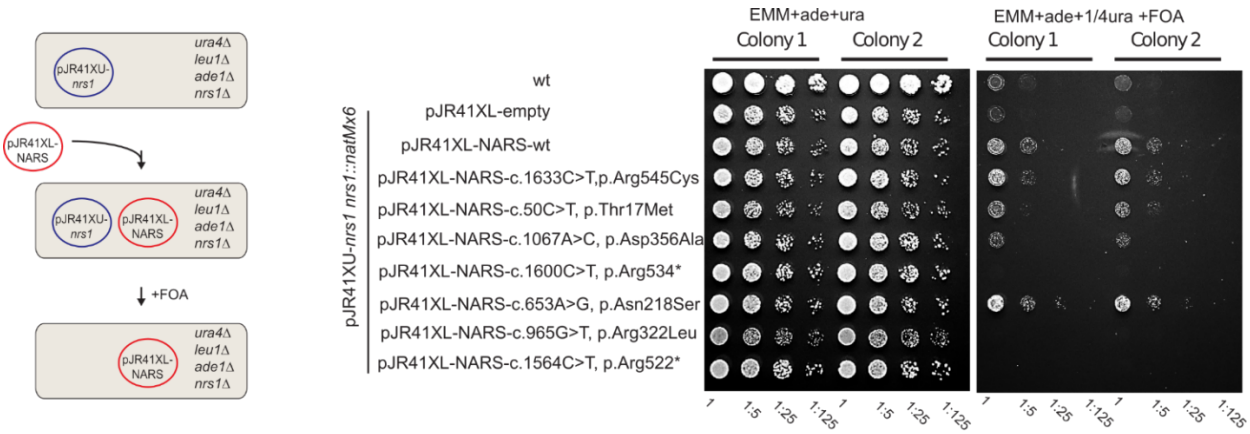


Figure 53. Human *NARS1* gene is able to complement fission yeast *nrs1*. Schematic of the yeast complementation assay. Fission yeast cells containing a plasmid expressing *nrs1* gene with uracil selectable marker whose genomic copy of the *nrs1* is deleted are transformed with plasmids containing the different variants of the human *NARS1* gene. These cells are promoted to lose the *nrs1*gene plasmids by incubating them in media with uracil for 24 hours. Five-fold serial dilutions of 2 different colonies of strains containing the different variants of *NARS1* gene wt, p.Arg545Cys, p.Thr17Met, p.Asp356Ala, p.Arg534*, p.Asn218Ser, p.Arg322Leu, p.Arg522*, the empty vector and wt cells were plated in media containing uracil, or media containing FOA, which allows the growth of only those cells that have lost the yeast *nrs1* plasmid and whose *NARS1* variant is able to complement *nrs1*D (Manole *et al.*, 2020, data generated by Maria Rodriguez Lopez).

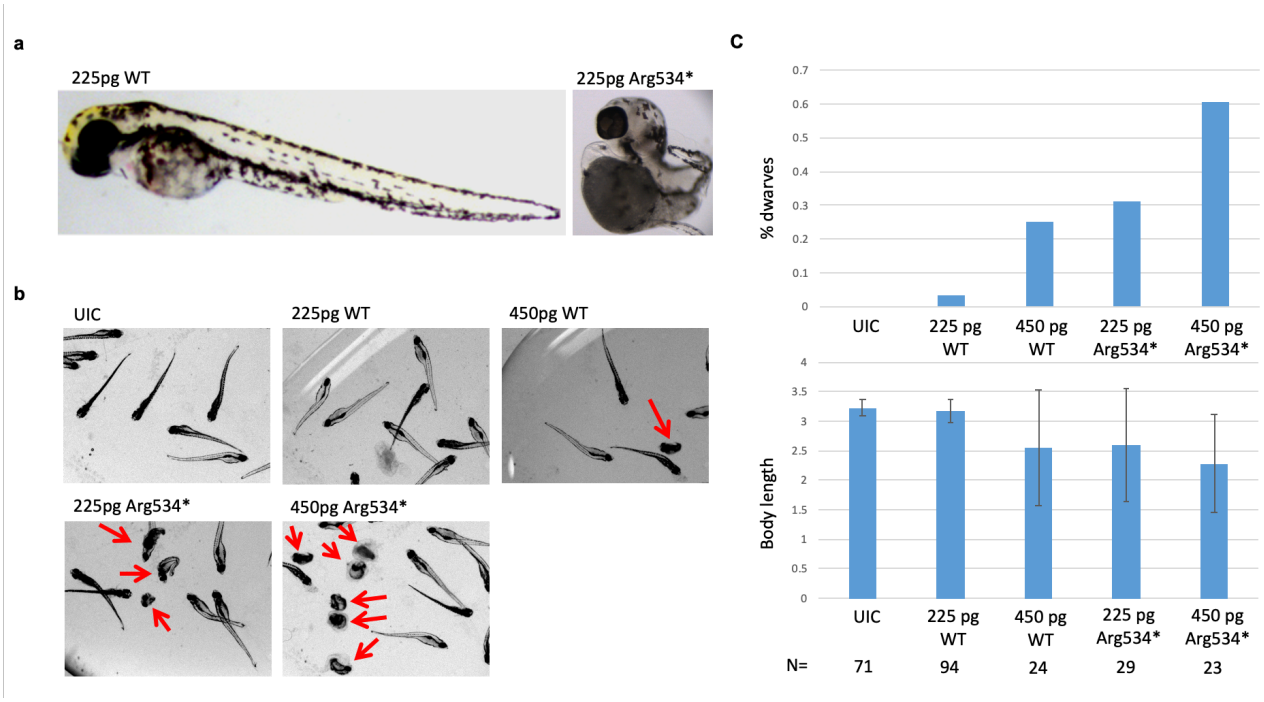


Figure 54. Microinjection of human *NARS1* RNA into zebrafish embryos. A) Microinjection of the indicated doses of wild-type (WT) or p.Arg534* mutant capped RNA encoding human *NARS1* into zebrafish embryos resulted in complete cyclopia and severe truncation of the body axis, a dwarf-like appearance. B) Low-power representative images of each experimental group at 5 days post-fertilisation. Dwarves are indicated by the red arrows. C) Quantification of animal body length and the proportion of animals exhibiting the dwarfic appearance for each experimental group, as indicated. Error bars are standard deviations (Manole *et al.*, 2020, data generated by Dagan Jenkins).

Given the essential function and constraint metrics of *NARS1*, in conjunction with the clinical phenotypes of included individuals, we propose that genotypes with dominant heterozygous variants produce a toxic gain-of-function. This is compared with the homozygous recessive variants that probably experience a loss-of-function, though this

can be least partially compensated for perhaps by other ARS genes. Taking into consideration the aminoacylation assay and yeast model for *de-novo* mutations, and the Western blot, aminoacylation assay and modelling for homozygous recessive mutations, we have confirmed the pathogenicity of all *NARS1* mutations mentioned. Similar effects are seen in *Aicardi*-Goutières syndrome which is caused by pathogenic variants in the *ADAR* gene. There are relatively frequent alleles that are pathogenic when *in trans* to a null, but never found in individuals with the homozygous state.[231] For distal C-terminus mutations, such as the homozygous p.Arg545Cys variant, the mechanism is likely due to abnormal protein structure and catalytic activity (Figures 43C, 45, 55, 56).

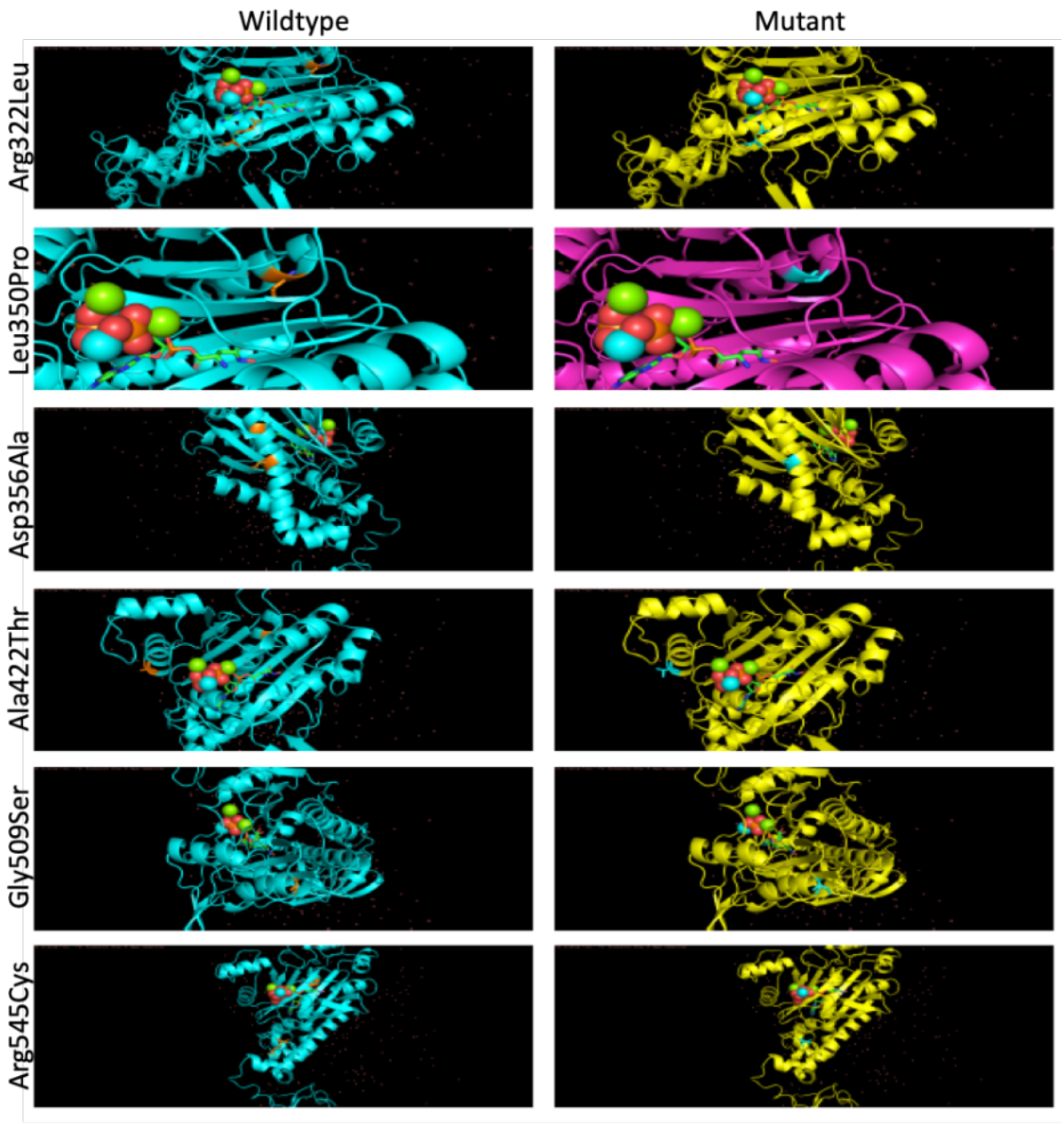


Figure 55. **Molecular modelling of *NARS1* mutations.** Wild-type NARS1 and each mutation was modelled using the Phyre^2 server[232]. We observed that each model gave rise to a "wild-type" AsnRS1 enzyme indicating that none of the mutations completely destabilizes the protein structure (left panels). The Arg11Pro, Thr17Met and Met34Leu mutations could not be modelled because they lie in the unknown UNE-N domain. The substrate depicted with balls on the figure corresponds to the ATP:Mg and L-Asp-beta-NOHandenylate:PPi:Mg from the structure of *B. malayi* NARS1 that was superimposed to each model we obtained. They were

superimposed with a very good concordance. From this study we propose that the mutations might have the following effects: Asn218Ser (affects the interaction between AsnRS1 and the anticodon arm of the tRNA), Arg322Leu (affects the stabilization of the aminoacyl-adenylate in the active site), Leu350Pro (little effect on AsnRS1 activity and dimer interface), Asn356Ala (might weaken the dimer stability), Ala422Thr (little effect on 3' tRNA end binding), Gly509Ser (slight interference on dimer interface), Arg545Cys (see Figure 45) (Manole *et al.*, 2020, data generated by Hubert Becker and Jean-Louis Mandel).

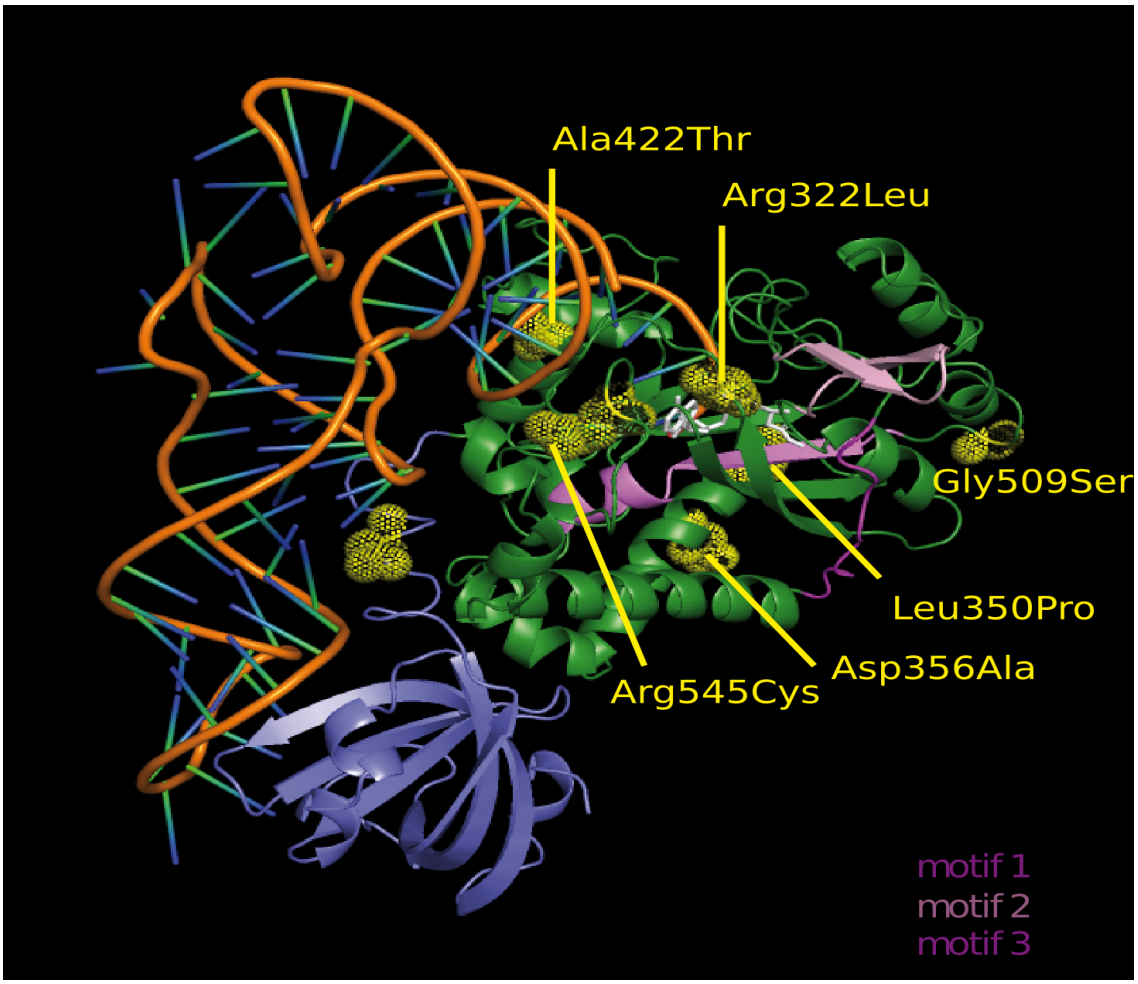


Figure 56. Model of human AsnRS1. This was obtained using the Phyre2 webserver. The latter was superposed with *Brugia malayi* AsnRS1 structure (2xti) complexed with ATP:Mg and L-Asp-ß-NOH adenylate:PPi:Mg (L-Asp-ß-NOH is

shown in white color on the figure). The tRNA is $tRNA^{Asp}$ that results from the superposition of the AspRS/$tRNA^{Asp}$ complex from yeast with the human AsnRS1 model. Note that the UNE-N domain is absent from the modelled structure. The position of the mutations are indicated by yellow spheres (Manole *et al.*, 2020, data generated by Hubert Becker and Jean-Louis Mandel).

Affected individuals had both central and peripheral nervous system involvement with a broad neurodevelopmental phenotype characterized by GDD, microcephaly, ataxia, neuropathy, and seizures. This is reflective of high *NARS1* expression in the cortex, cerebellum and brainstem as demonstrated in mouse brains (Figure 57).[233; 234]

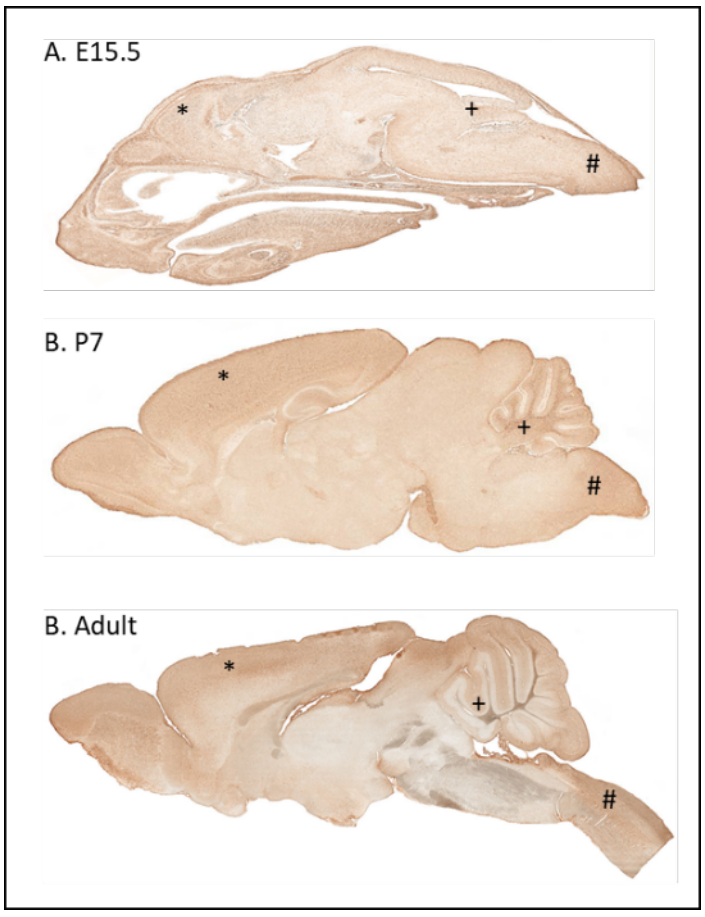


Figure 57. In-situ expression of the *NARS1* gene in mouse brain at three different ages, E15.5, P7 and adult. The *NARS1* probe used was GENSAT1-BX1431. Data

from the Gene Expression Nervous System Atlas (GENSAT) Project. Expression was moderate in the brain, higher in the cortex (*), cerebellum (+) and brainstem (#). Mutations have been reported for the majority of ARSs. AsnRS2, a mitochondrial ARS protein coded by *NARS2*, has recently been linked with an overlapping phenotype consisting of multisystem mitochondrial disorder (MID). Intellectual disability, epilepsy in childhood, hearing loss, and myopathy have also been seen in *NARS1* individuals,[235-238] In addition, aminoacyl-tRNA synthetase interacting multifunctional proteins 1-3 (AIMP1-3) participate together with nine cytosolic ARSs to constitute the so-called multi-synthetase complex, and have also been associated with a variety of human diseases. [239] With time, all ARSs are likely to have a disease association given their critical role.[213-215] The *NARS1* data brings the number of characterised ARSs to 35 out of 37. On a modified Taylor's Venn diagram of amino acid (AA) properties, *NARS1* is placed in close proximity to other amino acids with similar properties (*IARS1, LARS1, DARS1, EPRS1, NARS1, RARS1, QARS1*) which also have more severe phenotypes (Figure 58). [240]

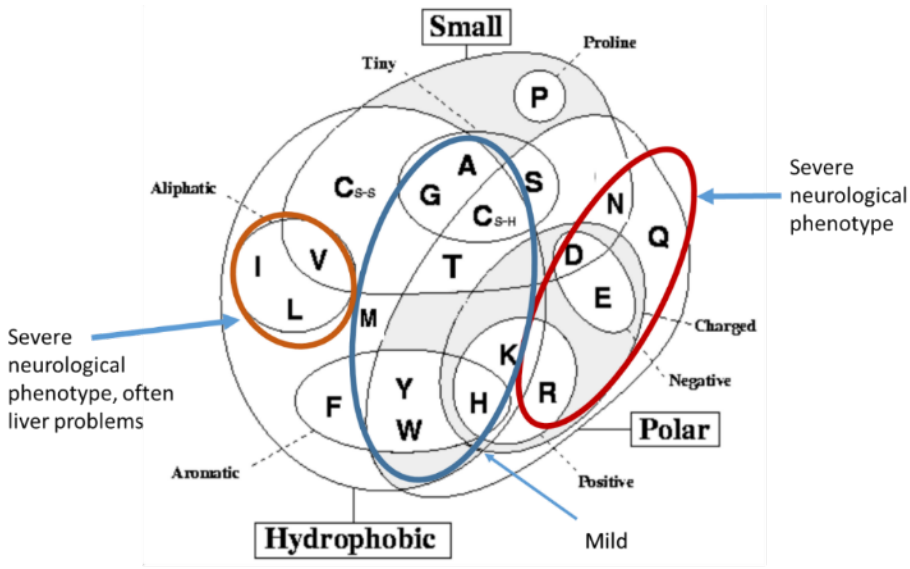


Figure 58. Modified Taylor's Venn diagram of amino acid properties, each amino acid is described by its physical or chemical properties[241]. AARs clinical phenotype

severity seem to cluster into two areas (red ovals) according to the amino acid that is targeted.

Our functional work supports a loss-of-function mechanism in homozygotes and has helped to further understand the role of *NARS1* mutations in disease. The development of CRISPR/Cas9 heterozygous knock-in and homozygous knock-out animal models is the next important step in understanding the molecular rationale of these *NARS1* variants. Considering the high number of individuals and variants identified here, the addition of *NARS1* to genetic testing panels for children and young adults presenting with NDD, epilepsy, and/or a demyelinating neuropathy may be of clinical benefit.

4.4.4 Further work. CRISPR-aided genome editing in mouse.

The generation of a point mutation within *NARS1* (the equivalent of human Arg545Cys is Arg556Cys) in mouse was reached by using CRISPR/cas9 technology. The experimental procedure followed is the same as outlined in Chapter 3. NFASC, in collaboration with the MRC Harwell team (Dr. Sara Wells, Dr. Lydia Teboul, Dr. Michelle Stewart) after successfully applying for their Genome Editing Mice for Medicine project.

Preliminary results for the Arg556Cys homozygous mouse show:

- Upon SHIRPA phenotyping done at 6 weeks of age, homozygous mice present with a lack of fluidity in movement on the gait test, an inability to trunk curl, 1 mouse falls through the bars on the jar positioning test, low scores on the locomotor test, and extended freeze time during the arousal test.
- Interestingly, the homozygous mutants seem to have a much higher activity, particularly in the dark phase. Upon videos processing and analysis of the

segments that fall over light/dark transition, we observed a peak in the mutant activity around those times (Figure 59).

- Almost all mice that were eventually sent to necropsy with abnormal gait turned out to have hydrocephalus. We were very surprised about this as we do see hydrocephalus in the background strain (in <1% of mice) and it's very obvious to us. Wild-type mice would normally get an obviously swollen head and in the early stages we can spot the swelling whilst the behaviour is still normal and the mouse appears otherwise healthy, however if it's developed more, the swelling becomes more pronounced and the mouse becomes inactive and hunched. This is not the series of events we have seen in the *NARS1* mice. For some of the *NARS1* mice that have hydrocephalus at necropsy, there is no evidence of that from looking at the mouse when it is alive.
- There is an obvious abnormal gait observed in the *NARS1* homozygous mice. This is also not something we have seen when wild-type mice develop hydrocephalus.
- There are a few other odd phenotypic observations as we have observed microopthalmia or anopthalmia in the NARS1 homozygotes, which combined with one mouse that needed teeth cutting (due to enlarged teeth) and several with hydrocephalus could be a result of a brain malformation syndrome.

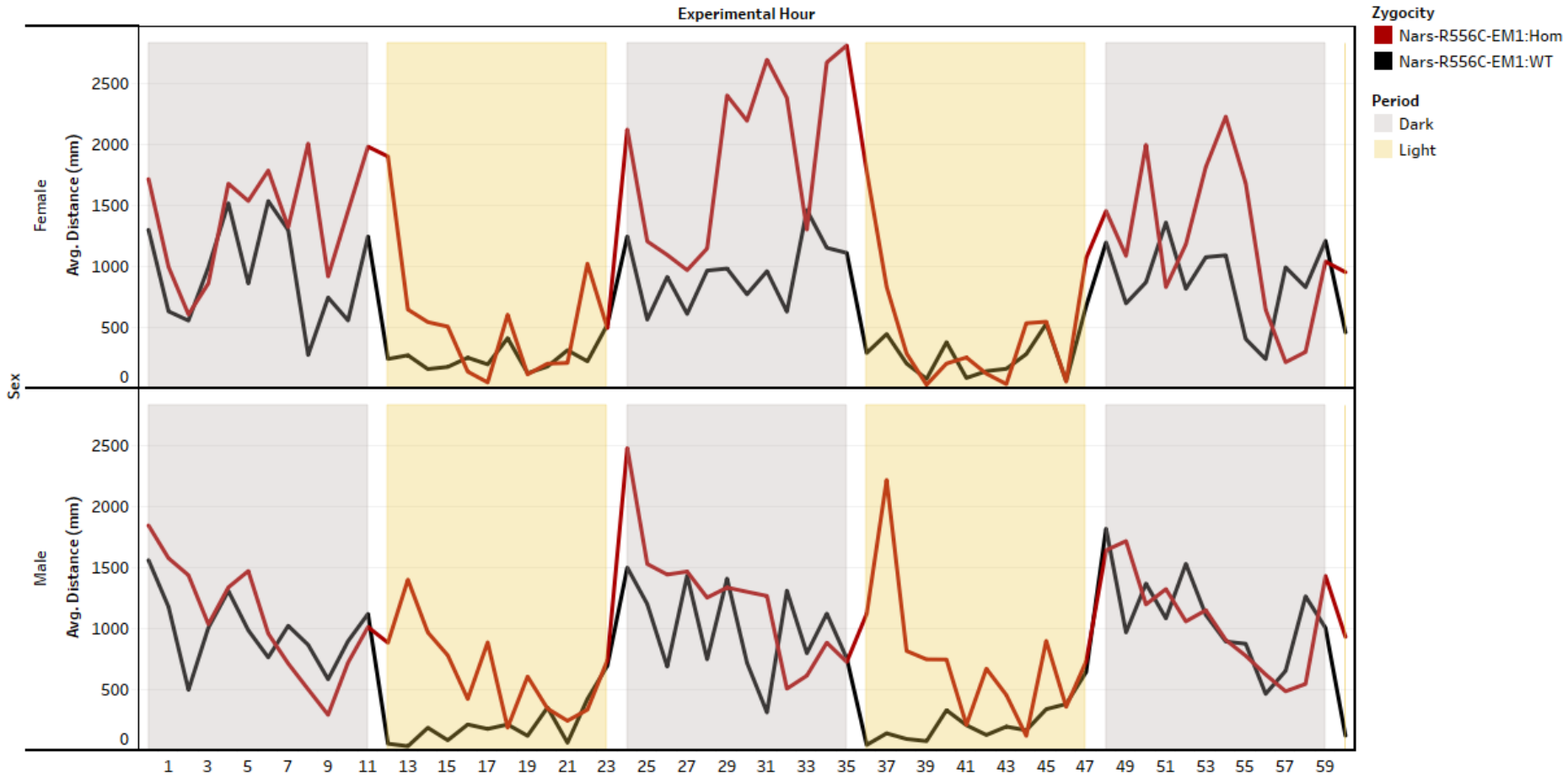


Figure 59. 8-week HCA data of the NARS-R556C-EM1-B6N-IC mice and the average distance travelled over 3 nights and 2 days. The Y-Axis represents the average distance travelled (mm) by each group, while the X-Axis represents the experimental hour, where experimental hour 0=19:00 day 0 and Experimental hour 60 = 7:00 day 3.

4.5 Expanding the phenotype and genotype spectrum of biallelic *TARS1* variants

4.5.1 Introduction

The human cytosolic threonyl-tRNA synthetase (TARS or ThrRS) is an homodimer (alpha2) which belongs to the aminoacyl-tRNA synthetase (ARS) protein family. It is a class II ARS protein that catalyzes the covalent binding of threonine to a threonine-specific tRNA (tRNAThr) to form an aminoacyl Thr-tRNAThr complex. It serves as an essential component of the translation machinery. TARS monomers need to form a dimer first in order to establish a catalytically active configuration that will allow specific binding of each monomer to the corresponding uncharged tRNA and free amino acid Thr.

The TARS protein is made up of several conserved domains which facilitate the amino acid charging (Figure 54B). These include the C-terminal anticodon binding domain, the catalytic core domain, a putative N-terminal editing domain, a so-called second additional domain (of antiparallel beta sheets) and a structurally uncharacterized TGS (for ThrRS, GTPase, and SpoT) domain important for nucleotide binding.

Biallelic variants in *TARS1* were first reported recently by Theil *et al*., 2019 in 2 unrelated patients presented with a brittle hair phenotype. They were found to carry compound heterozygous and homozygous missense *TARS1* variants, affecting the core catalytic domain. They functionally showed a profound effect on the TARS1

protein stability and enzymatic function, hereby expanding the spectrum of genes involved in trichothiodystrophy (TTD) to include genes implicated in amino acid charging of tRNA.[242]

Here, we present three novel threonyl-tRNA synthetase (*TARS1*) biallelic variants found by next-generation sequencing in a group of uncharacterized individuals presented with neonatal or infantile generalised tonic-clonic seizures, developmental delay including microcephaly and dysmoprhic features which to a degree resemble the thin brittle hair phenotype reported before. Homology models constructed with the guidance of Dr. Marie Sissler (University of Bordeaux) and enzymatic assays performed in collaboration with Dr. Marisa Mendes (Amsterdam UMC) revealed a significant decrease in TARS protein stability and enzymatic function. Future yeast complementation work done in collaboration with the Antonellis Lab (Michigan) will soon reveal whether the identified *TARS1* variants are pathogenic.

4.5.2 Results

Genetic findings

Biallelic *TARS1* variants were independently identified as the primary candidate genetic cause for the phenotype observed in the three families described herein (Figure 50 and Table 22). Homozygosity mapping carried out in Family 1 revealed a homozygous block in chromosome 5 and WES revealed the presence of two homozygous variants in *TARS1* (chr5:33453456 and chr5:33,459,830). These variants are predicted to cause a non-synonymous substitution (c.491G>A. p.Arg164His or c.1213G>A, p.Val405Ile) in the *TARS1* transcript NM_001258438.1. Sanger sequencing-based segregation analysis was performed in Family 1

confirming the presence of both variants. Further modelling and functional analysis, it is predicted to shed more light on which homozygous variant is the underlying cause of the patient's phenotype or whether both are contributing since we can't exclude this possibility.

Through GeneMatcher and the SYNaPS Study Group, we identified two additional cases, with clinical phenotypes partially overlapping that of Family 1. The exome enrichment evaluation for subject 2 showed 123x average coverage and at least 20x depth of coverage for 93% of the target. Among 74,395 high-quality variants, 13,008 affect either CDS or splice sites, whereas 435 have low or unknown frequencies according to the thresholds. A further variant stratification made by functional impact prediction score and considering a hypothesis of recessive condition allowed us to get a final set of 6 putative disease genes to be assessed for gene-phenotype correlations, two with X-linked transmission. The primary candidate gene was a homozygous variant in *TARS1* (chr5:33,467,055 G>A) predicted to cause a non-synonymous missense change (c.2087G>A, p.Arg696Gln) in the *TARS1* transcript NM_001258438.1 with CADD score of 34 and M-CAP score of 0.13.

WES carried out in Family 3 revealed compound heterozygous variants in *TARS1* (chr5:33,463,772 C>T and chr5:33,466,878 A>C) predicted to cause missense changes (c.1855C>T, p.Arg619Cys and c.1916A>C, p.Gln639Pro) in the *TARS1* transcript NM_001258438.1 with CADD scores of 23 and 22.7 respectively. Sanger sequencing-based segregation analysis was performed confirming that compound heterozygous variants c.1855C>T, p.Arg619Cys was inherited from the father and c.1916A>C, p.Gln639Pro from the mother.

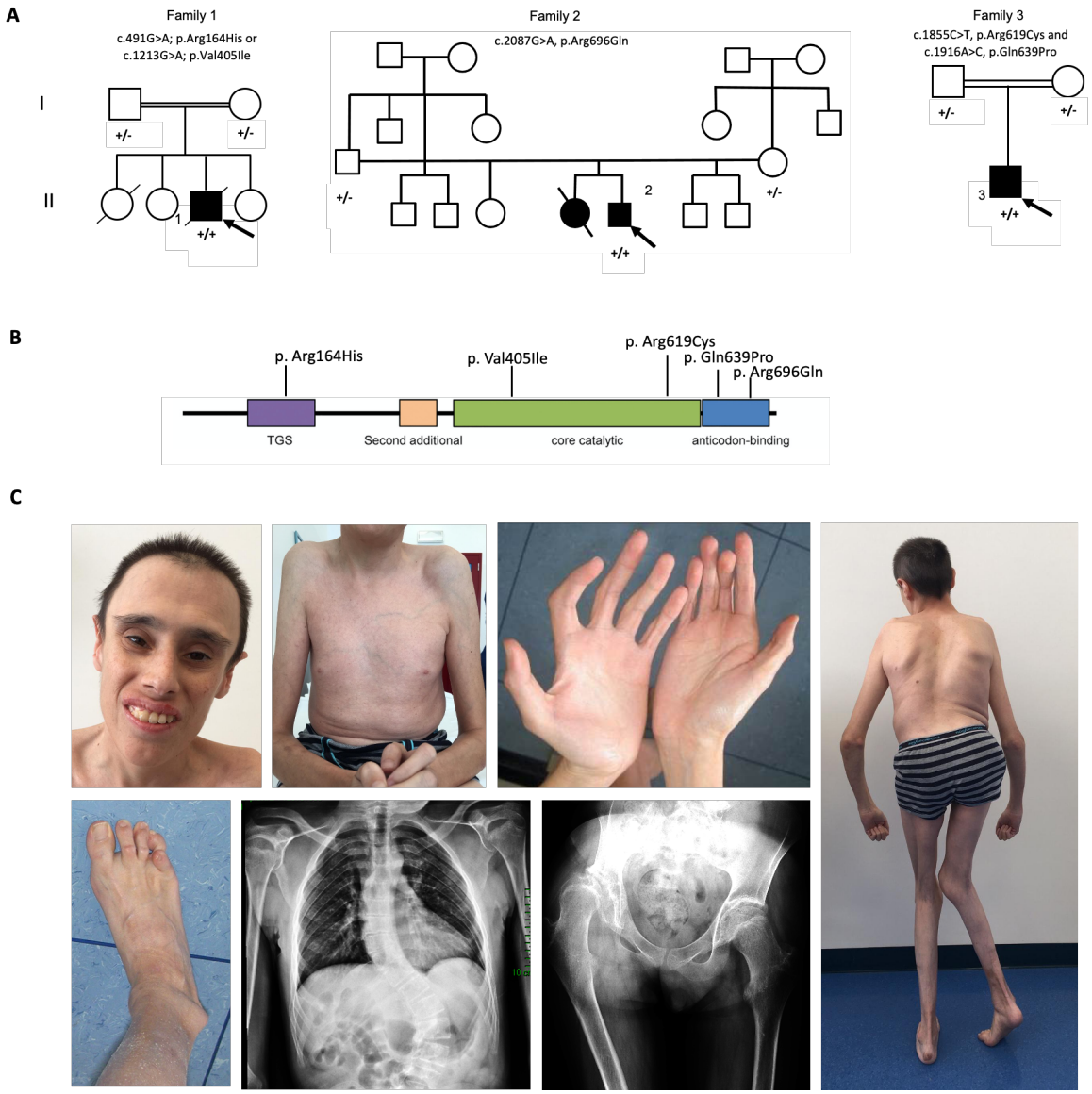


Figure 60. Clinicogenetic findings of individuals in our cohort. **A.** Pedigrees of the five families carrying biallelic *TARS1* variants. **B.** A schematic representation of the TARS1 protein showing the position of all *TARS1* variants identified. **C.** Clinical features of patient 2. Craniofacial dysmorphism: sparse hair, hypotelorism, upslanting palpebral fissures, beaked nose, small chin, large ears with prominent anthelix and absence of lobe. Skin assessment: thin transluscent skin, reduction in subcutaneous tissue. Other malformations: long fingers, 2nd and 3rd finger camptodactyly, rotational scoliosis and vertical talus. X-ray images: rotational scoliosis, hip dislocation and dysplastic acetabulum.

Clinical findings

The proband from family 1 is a 6-year-old male born to consanguineous Pakistani parents. He was born following a normal pregnancy and perinatal course, with microcephaly and low birth weight and height. In infancy, he failed to reach age-appropriate developmental milestones in terms of motor skills. Now aged 6, he has not acquired the ability to sit unaided, crawl or walk. In terms of speech, he shows no signs of language development to date and is unable to communicate. He began experiencing generalised tonic-clonic seizures in the neonatal period which persisted throughout childhood. His examination is notable for microcephaly with an OFC of 43 cm (<0.4th centile) and a rib cage deformity. He had severe truncal hypotonia with spasticity in the upper and lower limbs with brisk reflexes and ankle clonus. His coordination was also impaired. He was unable to co-operate with sensory examination. The clinical data was provided by Dr. Tipu Sultan.

The proband from family 2 is the second child born to Italian parents who are not known to be related. His older sibling was stillborn at full term with multiple malformations. He was born following an uneventful pregnancy at 40 weeks with a low birth weight (2nd centile, -2.3 SD) and microcephaly. At birth, he presented with agenesis of the left kidney, bilateral inguinal hernias, axial hypotonia and cryptorchidism due to gonadal dysgenesis which later required correction. Feeding difficulties occurred in the first two years of life and he exhibited signs of global developmental delay with failure to meet age-appropriate milestones across all adaptive domains. He experienced infantile febrile seizures and had a single generalized tonic clonic seizure, treatment with phenobarbital was therefore initiated and discontinued on resolution at age 9 years. He had moderate intellectual disability

but attended regular school with special support. His clinical course was complicated by several gastro-intestinal complications including gastroesophageal reflux disease, hiatus hernia, chronic gastritis (Helicobacter pylori-negative), chronic diarrhoea associated and small intestine bacterial overgrowth.

Now aged 23, he demonstrates several dysmorphic features on examination, including microcephaly OFC 50.5 cm (-4 SD), hypotelorism, upslanting palpebral fissures, beaked nose, small chin and large ears with prominent antihelix and absence of lobe. He has thin and fragile hair, thin, translucent and easily scratched skin with a reduction in subcutaneous fat. Multiple musculo-skeletal malformations include long fingers, 2nd and 3rd finger camptodactyly, rotational scoliosis and vertical talus. He is hypertonic with evidence of spasticity in his upper and lower limbs accompanied by brisk reflexes. He is also ataxic with poor limb coordination and dysmetria. Brain imagining detected multiple small calcifications located in the capsular nuclei and bilateral thalami, with cystic dilation of the “cavum vergae” and pineal gland. Additionally, X-ray imaging identified several skeletal abnormalities including severe kyphoscoliosis, hip subluxation with hypoplastic acetabular roof, small right pelvis, internal rotation of right leg and bilateral equinus of feet. The clinical data was provided by Dr. Marcello Niceta. The patient has provided informed consent for publication of the case.

The proband from family 3 is an Egyptian male born to non-consanguineous parents. He had an unremarkable perinatal course, with normal growth parameters. He presented with global developmental delay in childhood sitting for the first time at 5 years of age and now remains unable to walk unsupported aged 8 years old. His speech is also severely delayed with language development limited to single and

double syllables and lacking sentence formation. His examination is notable for craniofacial dysmorphisms including a long face with a pointed chin, flat narrow forehead, sparse eye brows, prominent nose, long philtrum, thin upper lip, mild prognathism and small low set ears. He has truncal hypotonia with limb spasticity and associated brisk reflexes particularly on the left. Eye movement examination revealed a gaze evoked nystagmus. His coordination was impaired evident by a bilateral intention tremor and dysmetria. Brain imaging was notable for cerebellar atrophy and a thin body of corpus callosum. The clinical data was provided by Dr. Maha Zaki.

Table 22. Summary of *TARS1* variants and clinical features of affected individuals

Variant	**c.491G>A; p.Arg164His c.1213G>A,p.Val405Ile**	**c.2087G>A, p.Arg696Gln**	**c.1855C>T; p.Arg619Cys, c.1916A>C; p.Gln639Pro**
Variant type	Homozygous	Homozygous	Compound Heterozygous
Inheritance	AR	AR	AR
Family	1	2	3
Individual(s)	1	2	3
Ethnicity/ country of origin	Pakistan	Italy	Egypt
Age at onset	Birth	Birth	Birth
Consanguinity	Yes	No	No
Presentation	Severe GDD	Moderate GDD	Severe GDD
ID	Severe	Moderate	Severe
Microcephaly	Yes	Yes	No
Dysmorphic	Yes	Yes	Yes
Seizures	Yes	Yes	No

Spasticity	Yes	Yes	Yes
Ataxia	Yes	Yes	Yes
AD = autosomal dominant, AR = autosomal recessive, GNDD = global development delay, ID = intellectual disability			

Analysis of aminoacylation activity

These experiments showed a reduction of *TARS1* enzymatic activity in patient-derived fibroblasts with *TARS1* variants. The most dramatic decrease was observed for patient 2 (carrying the p.Arg696Gln variant) (less than 5% when compared to healthy controls). The mildest decrease was observed for an additional candidate (patient 4 from the USA with compound heterozygous variants p.Asp115Asn and p.Lys636Arg) (~90% enzymatic activity compared to the controls) (Figure 61). This patient also had a pathogenic heterozygous variant in *YARS2* (c.201dupT, p.Gly68Trpfs*82), which fits better with the phenotype observed (progressive muscle weakness, limb girdle distribution, rapidly progressive peripheral muscle atrophy). Therefore, this patient was excluded from our cohort and we believe the *TARS1* compound heterozygous variants are not directly linked to his clinical picture. The fibroblasts from the Pakistani and Egyptian patients have not been tested yet but we plan to in the next months.

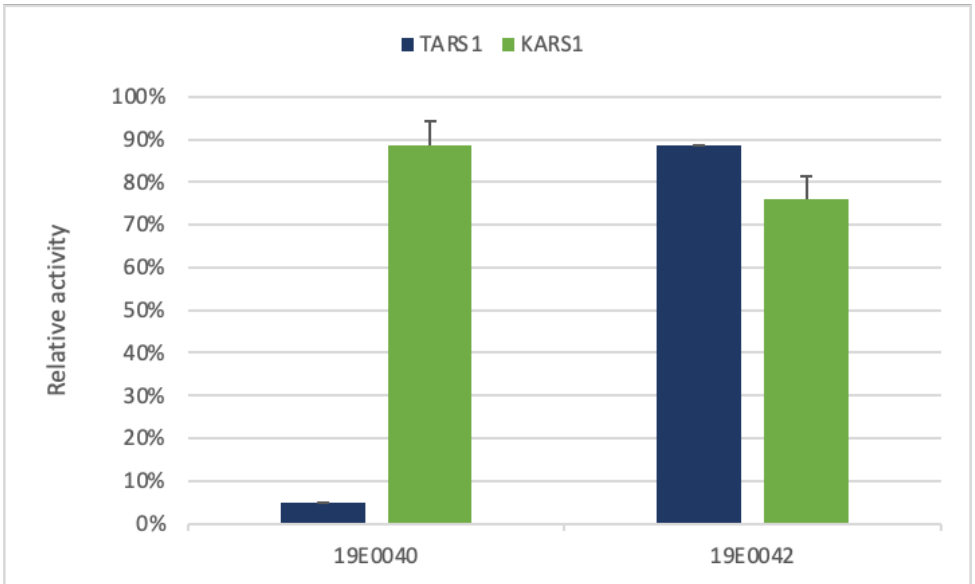


Figure 61. Amino acid charging on tRNA using patients' fibroblasts. The cytosolic fraction from patient 2 (id:19E0040) and 4 (id:19E0042), and two control fibroblasts was used for determining TARS1 (blue) and KARS1 (green; internal control) amino acid charging on tRNA. Threonine and lysine charging in control fibroblast C4RO was set at 100%, and error bars indicate SD of three independent experiments (data generated by Marisa Mendes).

Molecular modelling

Catalytic residues of TARS1 are less conserved and composed of three different motifs named motif 1, motif 2 and motif 3 (Figure 62).

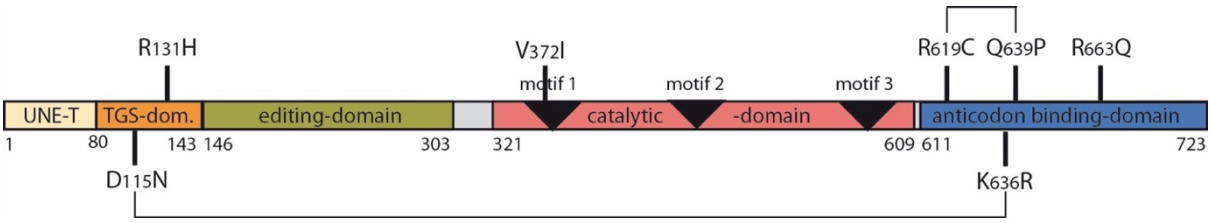


Figure 62. The modular organization of the TARS1 protein with all residues mentioned in this study and being tested. Note that residues shown on these structures are numbered according to NM_152295.5.

As seen in the cytosolic isoleucyl-tRNA synthetase (IleRS) protein model above, there is an editing domain. Some ARS may recognize wrong amino acids (when sizes and chemical properties are very close). In order to prevent that a given tRNA is improperly charged, an editing reaction occurs (in a dedicated structural domain) in order to ensure the fidelity of translation (by hydrolyzing the aminoacyl-tRNA bond when it is incorrect). To do this, the CCA extremity of the tRNA (3'-end) moves from the catalytic site to the editing site (for this quality control step) before being released.

The human threonyl-tRNA synthetase (ThrRS) has two additional N-terminal domains; the TGS and the UNE-T domains. The TGS is named after ThrRS, GTPase and SpoT proteins where it occurs. SpoT is also referred to a ppGpp hydrolase/synthetase. It is a ribosome associated protein that is activated during amino acid starvation. The precise function of the TGS domain remains unknown. The UNE-T domain, is a domain appended to the ThrRS, and was recently found to be implicated in the interaction of ribosome initiation factors, indicating that ThrRS could be implicated in translation initiation.[243] There is also evidence of another non-canonical function (beyond the aminoacylation property of the synthetase) of the ThrRS that would contribute it as an angiogenic marker.[244-246] Generally, the involvement of human cytosolic ARS in non-canonical functions is fairly widely documented,[247] which is not yet the case for mitochondrial ARSs.

Using the models explained further on, we have been able to represent ThrRS as a monomer, but *in vivo* ThrRS occurs as a dimer (Figure 63). In the case of compound

heterozygous mutations carried by patients identified in this study, we can anticipate (like for the mitochondrial AsnRS) that *in vivo* there is co-existence of three different proteins (two different homodimers and one heterodimer).

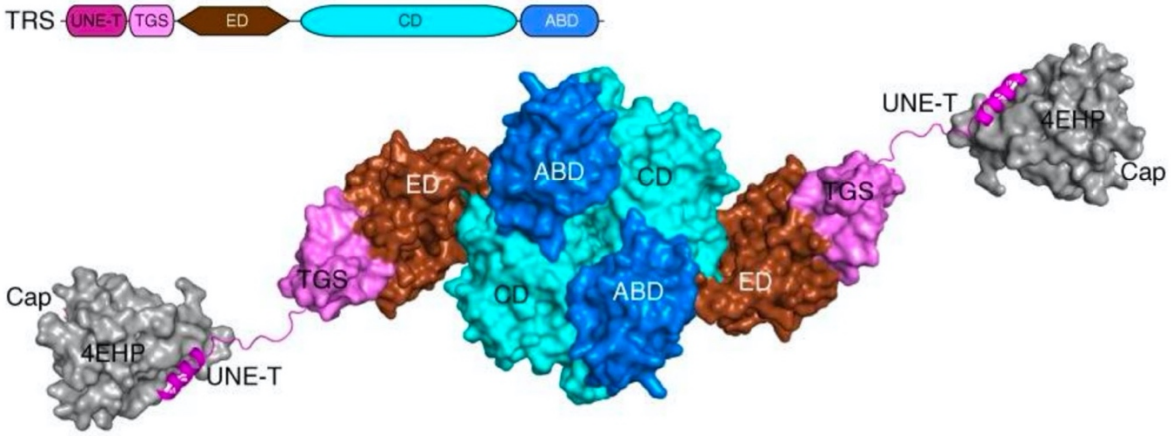
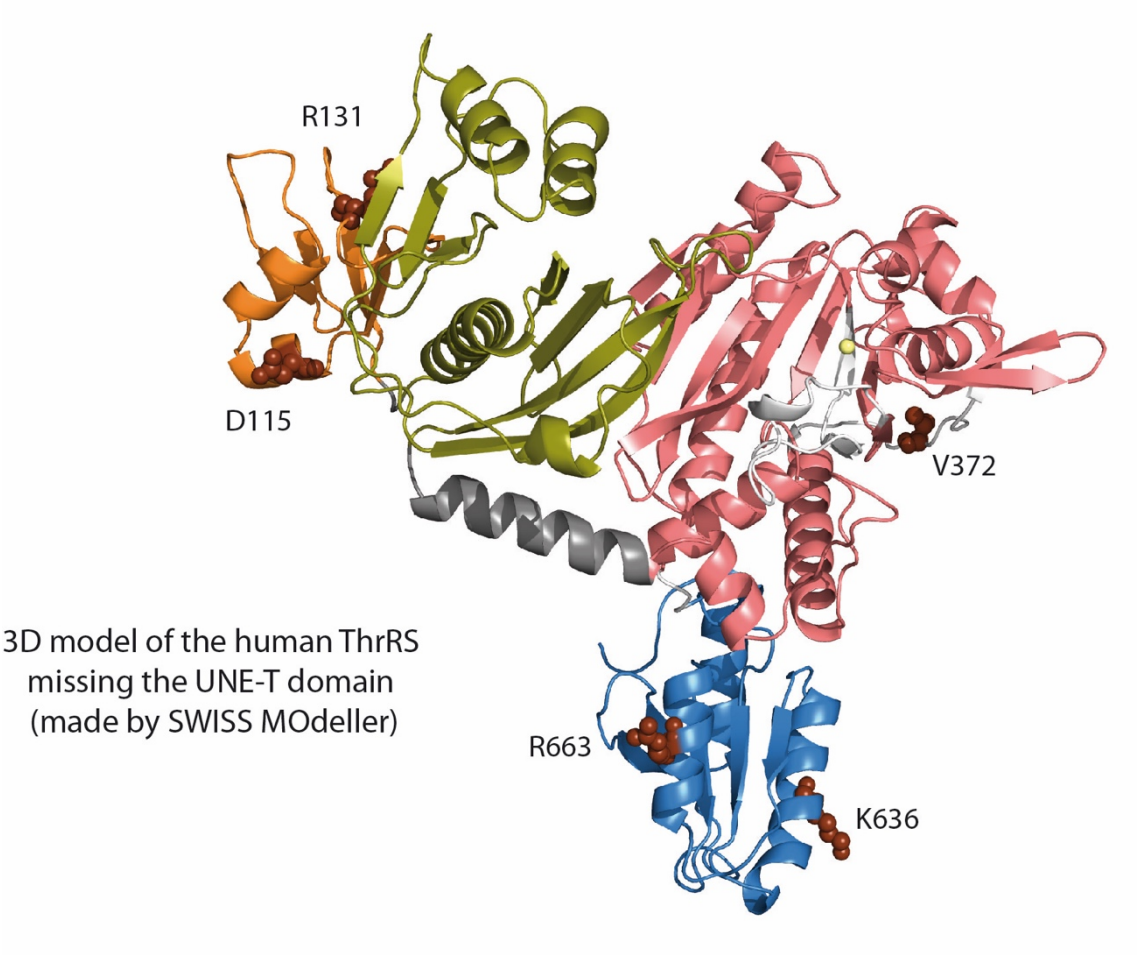


Figure 63. ThrRS dimerization showeing how the UNE-t domain interacts with the translation initiation factor. TGS: ThrRS, GTPase and SpoT, CD: C-terminal domain, ABD: anti-codon binding domain, ED: editing domain.

Analysis of the two homozygous variants identified in the Pakistani patient by amino-acid conservation within a robust multiple sequence alignment showed that Arg131 is highly conserved (>90%) but only in eukaryotes, and it is located within the TGS domain. A clear function of this extra domain has not been elucidated yet but it is thought to contribute to angiogenesis or modulation of the interaction of UNE-T with translation initiation factors. Val372 is located within motif 1 and is strictly conserved as Valine/Leucine/Isoleucine in the whole phylogeny, thus making it a silent mutation, most likely not disease-causing.

The canonical function of TARS1 in tRNA charging strongly depends on its structural conformation and correct assembly of a catalytically active dimeric TARS1 complex. The identified amino acid substitution p.Arg696Gln in subject 2, is also positioned at the anti-codon binding domain and is very likely to have an impact on the catalytic activity of the protein. We found it to be ~95% conserved in the whole phylogeny, while strictly conserved (100%) within eukaryotes. This residue is implicated in the tRNA anticodon loop recognition (see the bacterial structure in Figures 64 and 65) and makes a direct binding with residue 38 of the tRNA. This can clearly explain the drastic reduction in aminoacylation activity (<5%) seen in the Italian patient's fibroblasts. In theory, it is predicted that this variant probably impairs tRNA recognition.



3D model of the human ThrRS missing the UNE-T domain (made by SWISS MOdeller)

Figure 64. Crystallographic structure with the position of mutated residues mentioned so far. Note that residues shown on these structures are numbered according to NM_152295.5.

The *E. coli* structure is a meaningful and good model to be used in this analysis since it is a crystallographic structure that places the variants on the tRNA complex. Multiple sequence alignment was used to correlate the residues that are mutated in the human *TARS1* into the sequence from *E. coli* ThrRS.

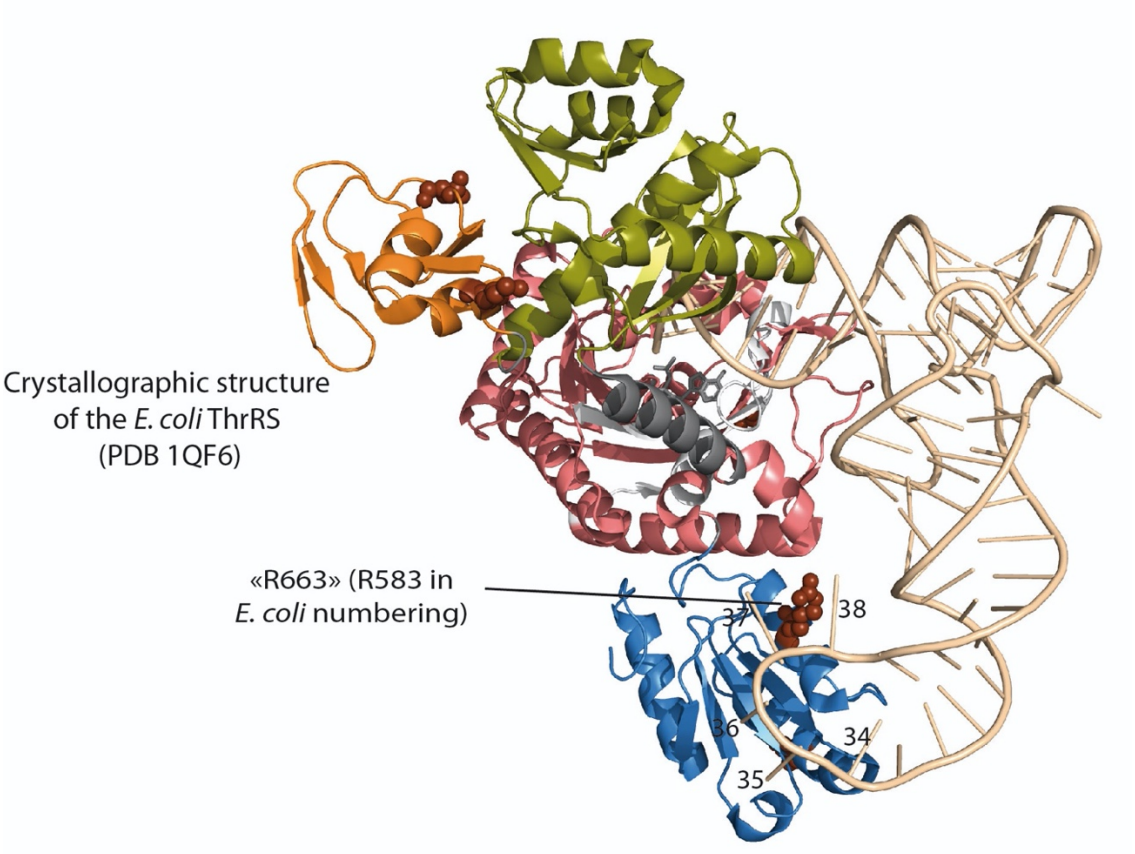


Figure 65. *E. coli* ThrRS crystallographic structure showing the position of the p.Arg696Gln of subject 2. Note that residues shown on these structures are numbered according to NM_152295.5.

Concerning the Egyptian patient, both variants (p.Arg619Cys and p.Gln639Pro) affect the anticodon binding domain (blue module), but in the opposite façade where tRNA is not interacting (see model with tRNA in Figure 66). A different model structure was used to see if the mutations can impact the dimerization of the enzyme.

The analysis of the amino-acid conservation within a robust multiple sequence alignment showed that Arg619 is conserved at >90% but only within the eukaryotes, while Gln639 is not conserved at all; even though replacement by a proline (a cyclic amino acid) might create a strong structural distortion.

To investigate if these two mutations can impair the dimerization of the ThrRS, the crystallographic structure of the dimeric human ThrRS that lacks the UNE-T and GST- domains (PDB 4TTV) was used. We docked on this structure the structure of the tRNA that was obtained in the crystallographic structure of *E. coli* ThrRS (PDB 1QF6). This allows to clearly situate the dimeric interface, and to situate the tRNA binding façade and the catalytic caveat (Figure 66).

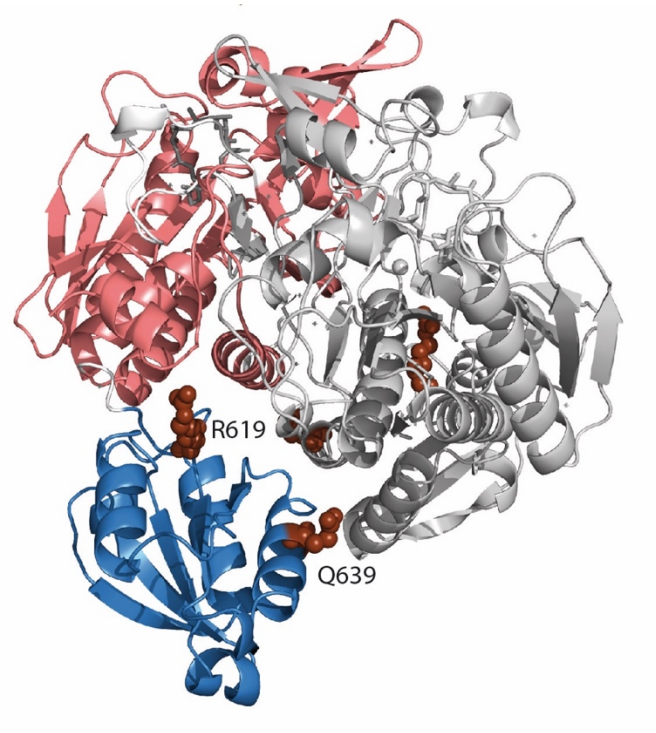


Figure 66. *E. coli* ThrRS crystallographic structure showing the position of the compound heterozygous variants p.Arg619Cys and p.Gln639Pro of subject 3. Note that residues shown on these structures are numbered according to NM_152295.5.

None of these two residues were found to be implicated in the dimerization, when using this modeling structure (Figure 67). Arg619 might contact Lys348 from the catalytic domain, which is in very close vicinity. Gly639 does not contact any residue but since it is mutated into a proline, this might again create a local distortion. This last model shows that none of the two residues are implicated in tRNA interaction.

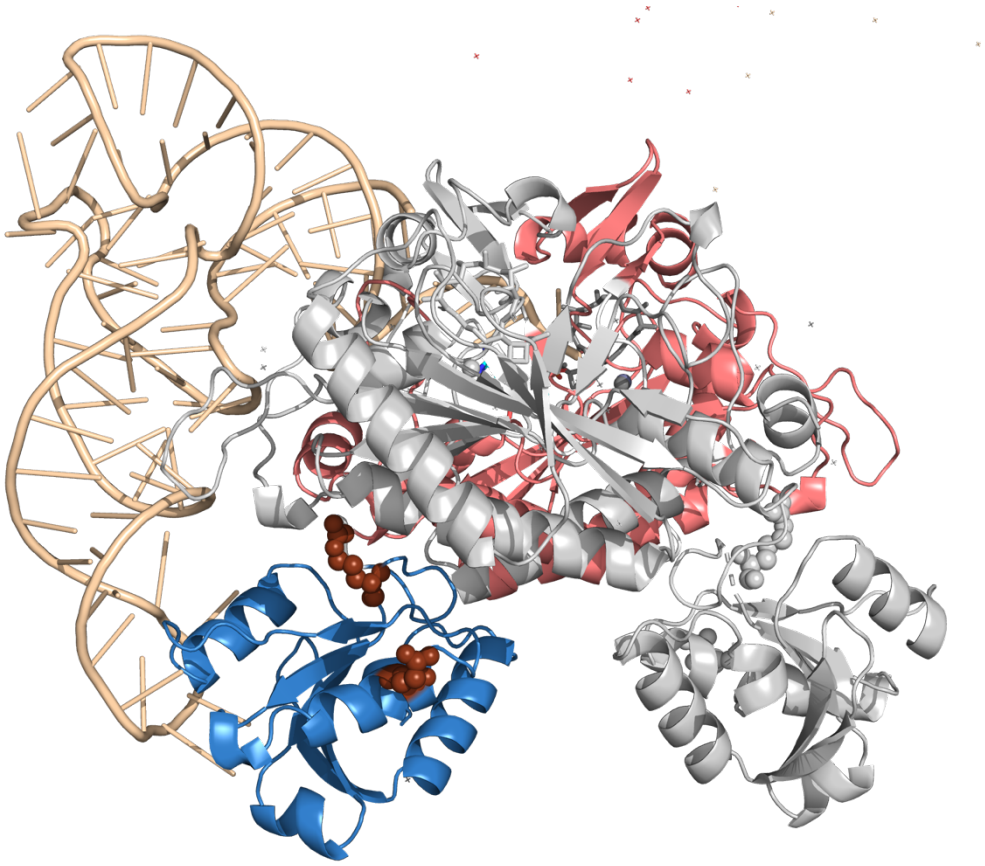


Figure 67. This model structure shows that none of the 2 compound heterozygous residues are implicated in tRNA interaction.

These modeling structures are a simplified representation and a theoretical analysis and can be used only to suggest some mechanistic possibilities. The two residues cannot be mutated at the same time in the same subunit to represent the patient's inherited variants which are found as compound heterozygous *in trans*. The effect of these variants was further investigated experimentally in patient-derived fibroblasts and aminoacylation activity was measured.

The identified amino acid substitution p.Asp115Asn in the American patient (not included in the current cohort of *TARS1* patients) is part of the protein domain, of

which the function remains unknown and might point to a non-canonical function. Recent crystallography work revealed novel strategies for providing specificity in tRNA selection. A novel protein fold, with a zinc ion found in the active site can make minor groove contacts with the tRNA acceptor stem and can therefore have a prominent role in tRNA identity and translational regulation (https://www.ncbi.nlm.nih.gov/pubmed/10319817). The second substitution p.Lys636Arg in the same subject is positioned at the anti-codon binding domain, where pathogenic variants do not always lead to decreased aminoacylation but the heterozygous state could account for the 20% reduction in activity. Nevertheless, aminoacylation activity was found to be not significant, ~90% compared to controls.

4.5.3 Discussion

TARS1 is already recently implicated in a rare recessive disease of non-photosensitive trichothiodystrophy (TTD). In this study, we present the molecular and clinical characteristics of three additional patients from three unrelated families with several unreported variants in *TARS1*. One patient had an overlapping phenotype with the two previously described cases but more importantly we have identified patients with early infantile epilepsy and co-occurring severe global developmental delay and dysmorphisms.

The enzymatic results presented so far but also the undergoing yeast complementation assays in collaboration with the Antonellis Lab (Michigan) and mRNA studies will soon hopefully shed light on the mechanism of disease and provide further support on the pathogenicity of the identified *TARS1* variants in the patient phenotypes.

We hope to soon be able to further expand the phenotypic and genotypic spectrum of the *TARS1*-related disorder. The list of ARS genes involved in neurodevelopmental disorders is increasing constantly to include genes implicated in amino acid charging of tRNA, important for protein translation. Further delineation of the molecular and phenotypic spectrum of *TARS1*-related disorders will help to improve management, develop treatments, and encourages the expansion of diagnostic genetic testing to include this gene.

Chapter 5. Epilepsies

Epilepsy is a neurological disorder characterized by abnormal synchronized electrical discharges in the brain due to symptomatic conditions such as infection, stroke, trauma, neoplasms, and autoimmunity or due to genetic predisposition. It is considered a common disorder across all age groups, with a prevalence of 3% during lifetime. The rapidity of disease gene discovery has resulted in remarkable advances in the field of epilepsy genetics (Figure 68).

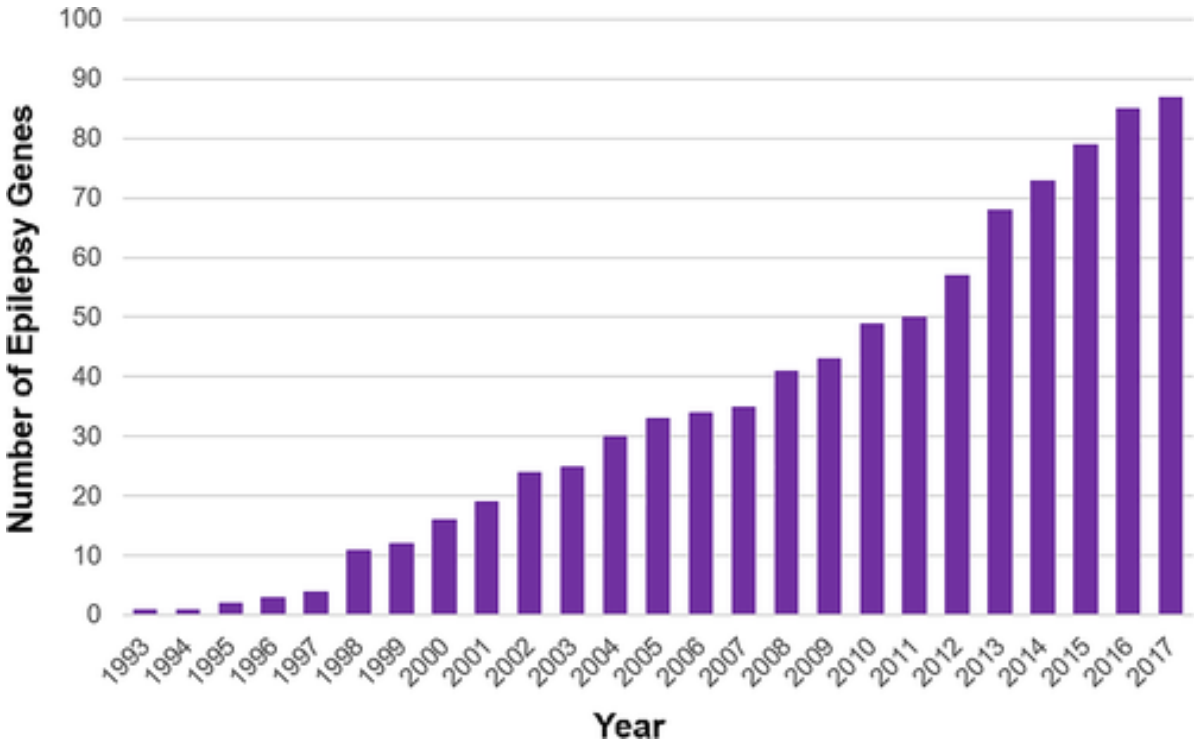


Figure 68. The cumulative number of known epilepsy genes by year. This does not include genes for disorders of metabolism and intellectual disability that may be associated with a high prevalence of epilepsy. [248]

Around 10.5 million children worldwide may suffer of epilepsy, representing approximately 25% of the global epilepsy population. Early onset, infantile epileptic encephalopathies are a heterogenous group of disorders describe a young

individual's cognitive and developmental profile as a direct result of the occurrence of seizures and the cerebral epileptic activity observed on electroencephalography (EEG). It is characterized by refractory to treatment seizures with an increased mortality rate. A significant degree of co-morbidity is associated with early infantile epileptic encephalopathy (EIEE) that includes global developmental delay, movement disorders, autism, and behavioural issues.

The cause of the encephalopathy is often genetic, with a recent scan of OMIM database showing a phenotypic series of 86 EIEE genes (April 2020). Whole exome or genome sequencing of the index patient and parents (trio-approach) provide a relatively quick molecular diagnosis with a high diagnostic rate. A rapid diagnosis is extremely important as a diagnosis can provide treatment options, such as GLUT1 deficiency or vitamin B6 dependent early onset epilepsy.[248]

Most neonate cases are caused by *de-novo* dominant mutations although X-linked and autosomal recessive forms are also frequent. Recently, our lab identified heterozygous de-novo *GRIA2* mutations in 28 unrelated patients with intellectual disability (ID) and neurodevelopmental abnormalities including autism spectrum disorder (ASD), Rett syndrome-like features, and seizures or developmental epileptic encephalopathy (DEE). [249]

5.1 Paroxysmal movement disorder and epilepsy caused by a *de-novo* truncating mutation in *KAT6A*

5.1.1 Introduction

KAT6A encodes for a histone acetyltransferase (HAT) prone to genetic and epigenetic modifications. Chromatin remodelling through histone acetyltransferase and deacetylase enzymes affects essential cellular processes. Performing whole exome sequencing (WES) with a trio-approach is extremely effective in identifying de-novo pathogenic variants. We report a *de-novo* truncating mutation in *KAT6A* (c.3338C>G; p.S1113Ter) in a young male presenting an exaggerated startle reflex and a complex neurological phenotype of autism, mental retardation and paroxysmal episodes. Nonsense mutations in genes involved in histone acetylation and deacetylation result in multiple congenital anomalies including intellectual disability.

5.1.2 Results

Clinical findings

A 21-year-old man, born to unrelated healthy parents of French and Scottish descent (Figure 69A), presented from birth with hypotonia, feeding difficulties and a small head circumference (31.75 cm; <3rd centile). Pregnancy and delivery were uncomplicated and there was no history of neurological disorders in the family, except for a maternal uncle who had epilepsy during his childhood. During the first months of life a delay of developmental milestones was evident and he also had infantile seizures, which spontaneously improved from the age of 3 years. Abnormal motor coordination and walking difficulties were evident in early infancy and

progressed over time. At the current age of 21, the patient has no autonomous gait, and requires a walker or sometimes a wheelchair. Speech was absent in his childhood and he still has no verbal communication. As part of his neurological phenotype, he has exaggerated startle responses to unexpected noises. These episodes usually last a few seconds with the patient experiencing generalized jerky movements that sometimes lead to sudden falls. The patient's paroxysmal episodes were not accompanied by any EEG ictal patterns, suggesting abnormal startle response rather than myoclonic seizures. Additionally, some distinctive facial features are also evident with a coarse face, dysplastic ears, prominent nasal bridge, prominent lower jaw, highly arched palate and mild bitemporal narrowing (Figure 69). Repeated electroencephalography and brain magnetic resonance imaging (MRI) studies performed during the follow-up were negative. Because of his complex clinical and neurological phenotype, he underwent several metabolic and genetic (including array comparative genome hybridization and panel sequencing of 55 genes causing intellectual disability) investigations that were all reported as normal. The clinical data was provided by Dr. Vincenzo Salpietro. The patient has provided informed consent for publication of the case.

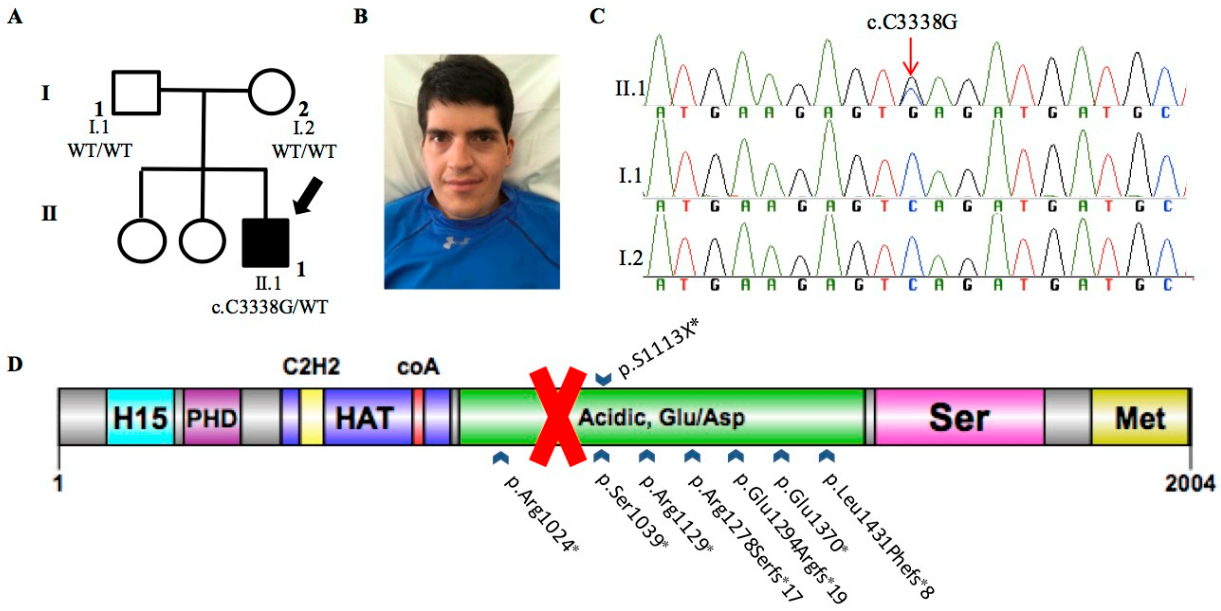


Figure 69. Clinicogenetic findings of the patient with the *KAT6A* variant. **A.** The pedigree diagram of the family. **B.** The 21-year-old proband affected by exaggerated violent startle reflex syndrome due to *de-novo KAT6A* mutation. Note the distinctive facial features of prominent lower jaw and highly arched palate. **C.** Individual results of Sanger sequencing showing the proband (II.1) carrying the *de-novo KAT6A* truncating mutation (c.C3338G; p.S1113Ter) which is absent in both the parents (I.1, I.2). **D.** A schematic illustration of the main functional domains of KAT6A protein: the nuclear localization domain (H15), a double-plant homeodomain finger (PHD), a histone-acetyl-transferase domain (HAT) containing a C2H2 zinc-finger domain (yellow), and an Acetyl-coenzyme-A-binding domain (red); an acidic glutamate/aspartate-rich domain and a transactivation domain with both a serine-rich (pink) and a methionine-rich (yellow) region. All previously reported KAT6A mutations indicated with arrows below while our mutated *de-novo* truncated version of KAT6A protein highlighted with the red cross and indicated with an arrow on the top.

Genetic findings

After following the filtering strategy used in chapter 4.2, we identified a *de-novo* non-sense variant in *KAT6A* (c.3338C>G; p.S1113Ter) as the most likely explanation for the disease pathogenesis. This is supported by a severe impact of the truncating mutation (Figure 69) and existing reports of patients (N=18) linking *de-novo* non-sense mutations in this gene to a similar phenotype (MIM# 616268) with global developmental delay, intellectual disability, poor growth, and absent or delayed speech. In addition, 3 out of 18 had feeding difficulties, 6 had facial dysmorphism and 4 had congenital cardiac defects (i.e. 3 an atrial septal defect, 1 a ventricular septal defect). [250-254] One of these cases had epilepsy that we observed in our patient, but none had the exaggerated startle responses. [252]

5.1.3 Discussion

The *KAT6A* gene (MIM# 601408; also known as *MOZ* or *MYST3) is* abundantly expressed in the brain, with the highest levels during fetal development, and encodes a histone acetyltransferase (HAT) component of a multi-protein complex involved in several cellular processes (e.g. differentiation, metabolism, apoptosis). [252; 255] The gene contains 18 exons and encodes a large 2004 amino acid protein, which includes a nuclear localization domain, a double C2H2 zinc finger domain that binds to acetylated histone tails, and a histone acetyl-transferase (HAT) domain (Figure 69D).

Of interest, *KAT6A* knock-out mouse models have been reported with severe abnormalities affecting the brain, gastrointestinal tract and skeleton.[256] Zebrafish genetic studies have also shown a crucial role of the chromatin remodeler gene

BRPF1 in regulating brain development through *KAT6A*- and *KAT6B*-mediated H3K23 acetylation. [255] Yan *et al.* [257] showed that non-sense mutations either in *KAT6A* or *KAT6B* cause deficiency of histone H3K23 acetylation, indicating a new emerging group of (clinically) overlapping developmental disorders due to abnormal histone H3 acetylation.

In conclusion, we describe a 21-year-old man with a complex clinical phenotype consisting in congenital microcephaly, hypotonia, global developmental delay, intellectual disability, seizures as an infant, absent speech, and a paroxysmal movement disorder consisting in exaggerated startle responses to sudden/unexpected noises. It is not clear if these are part of the underlying neurological disease or whether they represent functional (psychogenic) jerky movements, since we had not the opportunity to perform a neurophysiological evaluation to assess response and latency times. [258]

The neurological phenotype of *KAT6A*-related syndrome adding infantile epilepsy and exaggerated startle response as potential features of the clinical spectrum has been expanded. The future description of patients carrying *KAT6A de-novo* mutations will further clarify the neurological spectrum associated to this new syndrome.

5.2 Defining and expanding the phenotype of *PIGS*-associated early-onset epileptic developmental encephalopathy

5.2.1 Introduction

Congenital disorders of glycosylation (CDGs) are a rapidly growing heterogeneous group of genetic conditions. Inherited glycosylphosphatidylinositol-anchored protein (GPI-AP) deficiencies (IGDs) are a subset of the CDGs, which account for 0.15% of all developmental disorders.[259] In many cases, IGDs result from the failure of the GPI anchor to regulate anchored proteins (APs) on the external cell surface. This has consequences for early human development, neurogenesis, and fertilization, among others. [260-266] To date, 17 known IGDs lead to a wide range of symptoms including variable intellectual disabilities and developmental delays, seizures, hypotonia, weakness, ataxia, congenital malformations, and dysmorphic facial features. There are more than 150 human GPI-APs, often playing a role in CNS development, synaptic function, and plasticity. Several of which have been implicated in neurological disease.[260; 267]

Developmental and epileptic encephalopathies (DEE) include a range of severe epilepsies in which intractable seizures are accompanied by impairment of motor and cognitive functions.[268] Recently, many cases of IGD have been found among individuals with intellectual disability and intractable seizures. Early infantile epileptic encephalopathy (EIEE 55, [MIM] #617599), a severe form of epilepsy with frequent tonic seizures or spasms beginning in infancy, developmental delay or regression, has been linked to several genes important for GPI biosynthesis and attachment. For

example, *PIGB, PIGQ, PIGP,* and *PGAP1* have all been linked to novel autosomal recessive IGDs[269-271] while *PIGA* is an X-linked recessive disorder.[272]

Phosphatidylinositol glycan class S (*PIGS)* ([MIM] #610271) encodes a GPI-AP, specifically a subunit of the GPI transamidase complex that catalyzes the attachment of preformed GPI to proteins containing a C-terminal GPI attachment signal.[273] *PIGS* has only recently been linked to human disease in seven individuals from four unrelated families with early infantile epileptic encephalopathy (EIEE), all with compound heterozygous or homozygous variants in this gene.[274; 275] Phenotypes included severe global developmental delay, seizures, hypotonia, weakness, ataxia, and dysmorphic facial features, but also multiple joint contractures consistent with fetal akinesia in two fetuses. All seven individuals showed a GPI-AP deficiency profile. Notably, one individual showed a reduction in seizure frequency with pyridoxine (100 mg per day).

In this study, we present the molecular and clinical characteristics of six additional affected individuals from five families with several unreported variants in the *PIGS* gene. All patients have intractable early infantile epilepsy and co-occurring severe global developmental delay. All individuals have brain malformations of varying degree, hypotonia, and microcephaly. Other findings include vision impairment, hearing loss, renal malformation, and hypotonic facial appearances with minor dysmorphic features without a distinctive facial gestalt. Four individuals died of neurologic complications. GPI anchoring studies performed in the Campeau Lab (Montreal) on one individual revealed a significant decrease in GPI-APs. We confirm that biallelic variants in *PIGS* cause vitamin pyridoxine-responsive epilepsy due to

inherited GPI deficiency and expand the phenotype and genotype of *PIGS*-related disorder. Further delineation of the molecular and phenotypic spectrum of *PIGS*-related disorders helps to improve management, develop treatments, and encourages the expansion of diagnostic genetic testing to include this gene.

5.2.2 Results

Clinical findings

We identified six novel patients from five families with biallelic variants in the *PIGS* gene identified by trio exome sequencing (Figure 70). Parental consanguinity was reported in four of the five families, all of whom had homozygous variants in *PIGS*. The clinical features of the affected individuals are summarized in Supplementary Table 9. All individuals presented with intractable early infantile-onset seizures, severe global developmental delay, and delayed speech. Patients 2, 3, and 5 died from complications of their neurological disease (recurrent respiratory infections). Intractable seizures requiring multiple antiepileptic drugs (AEDs) were required in all patients. The EEG patterns in our cohort was variable, influenced by the age of the patients and the effect of the specific variant but all the subjects showed EEG recording indicative of a severe developmental epileptic encephalopathy, namely, pseudo-hypsarrhythmia with runs of high amplitude and slow waves predominant over posterior regions associated, multifocal epileptic abnormalities within a diffusely slowed high-amplitude during wakefulness and sleep or within a burst-suppression pattern. All of the individuals in this cohort had abnormal findings on brain MRI, including atrophy of the frontal and anterior temporal lobes, with or without generalised brain atrophy. Frontal atrophy led to the development of bilateral subdural hygromas in one patient (patient 4). We noted hypoplasia of the pons in two

patients (patients 3, 6). One individual (patient 6) with severe pontine hypoplasia has large thalami and massa intermedia. All six participants have microcephaly (Figure 71). Behavioral anomalies in two siblings (Patients 4 and 5) include autistic features and excessive laughing (Patient 4) and excessive crying (Patient 5). Variable other features included renal malformation (1/6, 17%), vision impairment (3/6, 50%), severe hearing loss (2/6, 33%), and acquired arthrogryposis (2/6, 33%). Facial dysmorphism was observed in most of the affected individuals, mainly consisting of coarse features such as thick arched eyebrows, almond-shaped palpebral fissures, depressed nasal bridge, and deep philtrum.

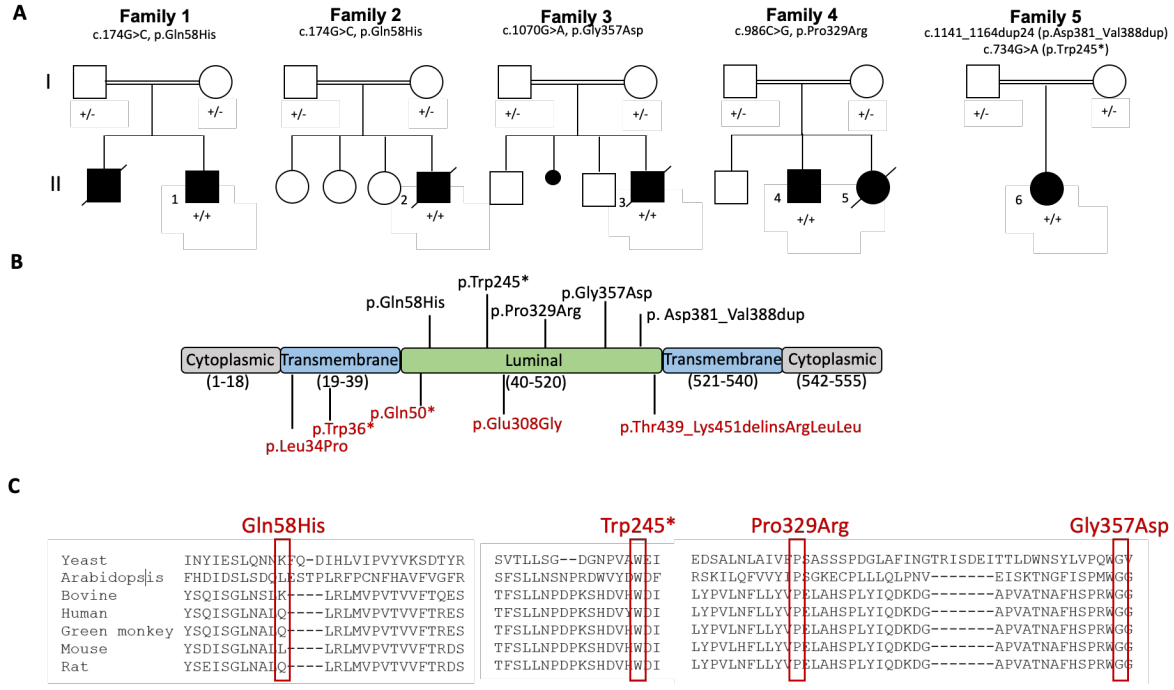


Figure 70. Genetic and molecular findings of *PIGS* patients. **A.** Pedigrees and segregation results (+ represents the variant) of the five families carrying biallelic *PIGS* variants. **B.** A schematic representation of the PIGS protein showing the position of all *PIGS* variants identified, with previously reported variants in red. **C.** Inter-species alignment performed with Clustal Omega shows the complete

conservation down to invertebrates of the amino acid residues affected by the substitutions.

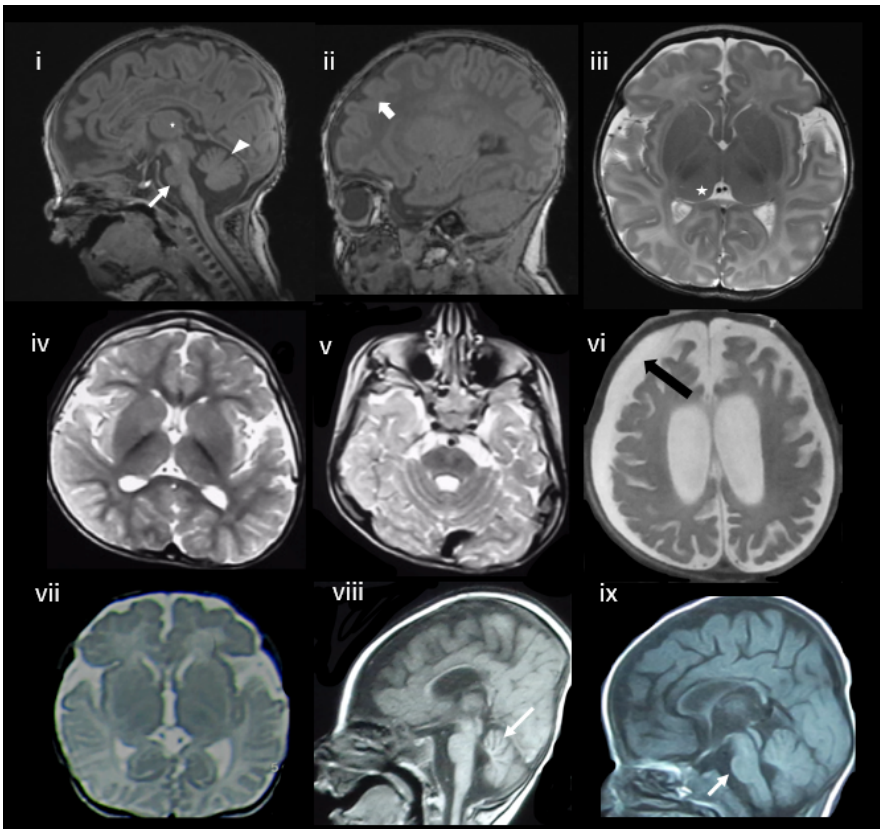


Figure 71. Radiological findings of *PIGS* patients. MRI findings of **i-iii**) patient 6 at 5 weeks. Sagittal T-weighted image at the midline (i) shows small pons (arrow), early superior cerebellar vermian atrophy (arrow head) and a large massa intermedia (star). Parasagittal T1-weighted image (ii) shows underdeveloped and undersulcated frontal lobes (arrow). Compare the sulcation in the anterior aspects with the posterior cerebral hemisphere. Axial T2-weighted image (iii) shows under-opercularization due to underdevelopment of the frontal and anterior temporal lobes and large thalami (star). iv-v) patient 4 at 5 years. Axial T2-weighted images at the supratentorial (iv) and infratentorial (v) levels shows generalised brain atrophy and delayed maturation of myelin. Note again the worse atrophy anteriorly in the frontotemporal regions.vi) patient 5 at 5 months. Axial T2-weighted image showing atrophy related bilateral

subdural hygromas (arrow). vii) patient 1 at 3.5-months. Axial T2-weighted image. Under-opercularization shown bilaterally secondary to fronto-temporal atrophy. Frontal dysgyria is also noted. viii) patient 2 at 34 weeks. Sagittal T1-weighted image shows microcephaly, with greater volume reduction in the frontal lobe, as well as early superior cerebellar vermian atrophy (arrow). ix): patient 3 at 14 months. Sagittal T1-weighted image shows a small pons and globally reduced cerebral white matter volume (worse anteriorly) associated with hypoplasia of the corpus callosum.

Though seizure management with AEDs is challenging, there is potential improvement related to pyridoxine and folinic acid administration. Patients 1 and 6 were both on pyridoxine (100 mg per day) and folinic acid (1 mg/kg/day per day) and have reported improvement in neurologic function; Patient 1 reported decrease in seizure frequency and Patient 6 reported improved developmental progress.

In family 1, a 3-year-old male was born preterm at 7 months gestational age to consanguineous Pakistani parents. He presented with neonatal-onset multifocal clonic and myoclonic seizures, global developmental delay, impaired vision, sensorineural hearing loss, and stereotypic mouth twitching. Electroencephalography revealed pseudo-hypsarrhythmic pattern during wakefulness with runs of 2-3-Hz high amplitude and synchronous slow waves predominant over posterior regions associated with few spikes and slow-wave spikes over temporal regions (Figure 72). His seizures have been managed by various AEDs, including levetiracetam, nitrazepam, and topiramate. Supplementation with pyridoxine (100 mg per day) reduced the frequency of seizures for the first two years. At 3 years old he was microcephalic at 43cm (<0.4th centile) with no growth failure or malformations. Brain

MRI done at 3.5 months of age showed under-opercularization bilaterally secondary to frontotemporal atrophy, as well as frontal dysgyria (Figure 71). He had an affected brother who started to have focal-clonic seizures since the first month of life. He was developmentally delayed with no gaze fixation or neck holding. EEG was suggestive of epileptic encephalopathy and he died at 4 months of age due to respiratory infection. The clinical data for families 1, 2 and 3 was provided by Dr. Tipu Sultan.

In family 2, a male was born from consanguineous parents from Pakistan following an uneventful pregnancy. He did not cry at birth and later presented with neonatal multifocal clonic seizures, developmental delay, as well as vision and hearing impairment. His EEG at age 1.5 years revealed multifocal epileptic abnormalities within a diffusely slowed high-amplitude during wakefulness and sleep intermixed with a burst-suppression pattern, indicative of developmental epileptic encephalopathy (Figure 72). He was treated with multiple antiseizure medications but not pyridoxine supplementation. He was microcephalic with no significant growth failure (unknown parameters) and behavioral abnormalities. Brain MRI done at 1 month of age revealed microcephaly, with greater volume reduction in the frontal lobe, as well as early superior cerebellar vermian atrophy (Figure 71). He died at the age of 1 year 9 months status epilepticus.

In family 3, a male was born after an uneventful pregnancy and delivery to consanguineous parents of Pakistani origin. He had developmental delay and febrile generalized tonic-clonic seizures. Seizures began at 2 months of age, and throughout infancy, multiple seizure types developed. He started with focal-clonic seizures which became multifocal by 8-10 months of life followed by tonic spasms in

the late first year with progression to generalized tonic-clonic seizures in the second year of life. EEG at age 3.5 years showed multifocal epileptic abnormalities within a burst-suppression pattern (Figure 72). He partially responded to a combination of levetiracetam, carbamazepine and valproic acid and he has not been treated with pyridoxine supplementation. He was non-verbal and microcephalic (unknown parameters) with no behavioral symptoms. His neuromuscular examination revealed spastic gait and truncal hypotonia. His brain MRI done at 14 months revealed a small pons and globally reduced cerebral white matter volume associated with hypoplasia of the corpus callosum (Figure 71). He died at the age of 2 years due to respiratory failure following an acquired respiratory tract infection.

In family 4, two affected siblings (Patient 4 and Patient 5) were born to consanguineous Egyptian parents after an unremarkable pregnancy. Patient 4 was an 8-year-old male with severe global developmental delay, neonatal-onset intractable tonic seizures (GTS), mild autistic features, and excessive laughing. From the age of 12 months, he suffered from recurrent seizures with focal onset followed by tonic-clonic seizures, triggered by fever. He has not been responsive to a combination of valproate, levetiracetam, topiramate, and clonazepam resulting in frequent admissions to the intensive care unit. At age 8, his growth parameters included weight at 16kg (<0.4th centile) and height at 115cm (<0.4th centile). He was microcephalic (46.5 cm (<0.4th centile). He was non-ambulatory and non-verbal. He had mild hypotonia (able to sit and stand supported), brisk reflexes, and mild acquired arthrogryposis of the knees. Brain MRI showed generalised brain atrophy (worse anteriorly in the frontotemporal regions) and delayed maturation of myelin (Figure 71). At age 6 years, his EEG showed diffusely slowed, low-voltage,

monomorphic background activity indicative of subcortical -cortical dysregulation with no evidence of epileptic abnormalities (Figure 72). The affected sister (Patient 5) similarly had global developmental delay with neonatal onset of intractable focal seizures followed by secondary generalized myoclonic jerks associated with lip-smacking and eye blinking movements. Her seizures were not responsive to valproate, levetiracetam, topiramate and phenytoin. At 9 months of age, she was not alert, did not follow objects, had no acquired milestones, and is fed by nasogastric tube. Growth parameters at 9 months include weight at 7.7kg (1^{st} percentile) and height at 65cm (5^{th} percentile). She has microcephaly at 40cm (94^{th} percentile). She presents with excessive crying, hypotonia, and spasticity. Her brain MRI shows atrophy related bilateral subdural hygromas (Figure 71). She died at the age of 1 due to intractable seizures after being ventilated for 2 weeks. Neither sibling was treated with pyridoxine supplementation. The clinical data for family 4 was provided by Dr. Maha Zaki.

In family 5, patient 6 is a 2-year-old female born to non-consanguineous parents of Chinese descent. She presented with multifocal seizures at four weeks of life. EEGs showed frequent independent and synchronous spike-wave discharges over right posterior quadrant and occasionally at left posterior quadrant. Her seizures were not well controlled on multiple AEDs. Her current treatment includes clobazam, cannabidiol, diazepam, folinic acid, levetiracetam, perampanel, pyridoxine, zonisamide, and clonazepam. Despite treatment, at 24 months her seizures occur multiple times daily, often with rapid eye blinking and lip-smacking, consistent with myoclonic seizures. In addition to her seizures, she currently presents with global developmental delay, cortical visual impairment, and profound hypotonia. Though

seizures have not been well controlled, parents report an increase in eye tracking, alertness, and fine motor skills after starting B6. Her development has slowly improved and at 16 months she has a social smile and is making vocalizations, though she does not babble or have any words. She rolls but cannot sit without support or support her head consistently. Growth parameters at 24 months include weight of 9.32 kg (5^{th} centile), length of 83 cm (24^{th} percentile). She was normocephalic at birth then developed progressive microcephaly with head circumference of 43.6 cm (<0.4^{th} centile). She only eats pureed foods due to poor swallow coordination. Brain MRI showed small pons (Figure 71, arrow), early superior cerebellar vermian atrophy (arrow head) and large massa intermedia (star), as well as underdeveloped and simplified frontal lobes (arrow) and opercularization due to underdevelopment of the frontal and anterior temporal lobes and large thalami (star). Following a febrile urinary tract infection, renal ultrasound identified asymmetric kidneys with her right kidney being 1.4 cm larger than the left and possible duplicated collecting system of the right kidney. The clinical data for family 5 was provided by Dr. Bianca Russell.

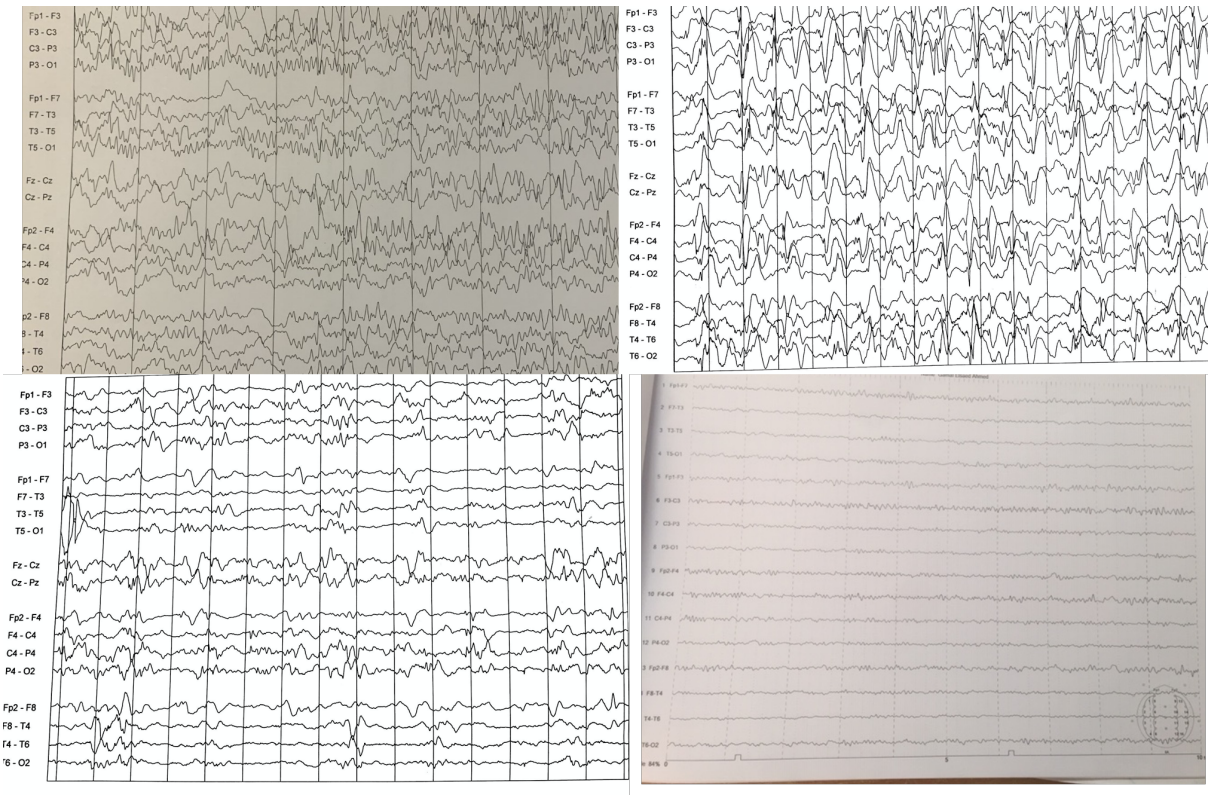


Figure 72. EEG recordings of patients harbouring *PIGS* variants; patient 1 (top left) performed at age 3 years showing pseudo-hypsarrhythmic pattern during wakefulness with runs of 2-3-Hz high amplitude and synchronous slow waves predominant over posterior regions associated with few spikes and slow-wave spikes over temporal regions, patient 2 (top right) done at age 1.5 years showing multifocal epileptic abnormalities within a diffusely slowed high-amplitude during wakefulness and sleep intermixed with a burst-suppression pattern, indicative of active epileptic developmental encephalopathy, patient 3 (bottom left) done at age 3.5 years shows multifocal epileptic abnormalities within a burst-suppression pattern compatible with active epileptic developmental encephalopathy, patient 4 (bottom right) done at age 6 years shows diffusely slowed, low-voltage, monomorphic background activity indicative of subcortical -cortical dysregulation with no evidence of epileptic/paroxysmal abnormalities.

Genetic findings

Exome sequencing in the probands of families 1 and 2 revealed a novel homozygous variant in exon 2 of *PIGS*, c.174G>C (p.Gln58His) predicted to affect splicing (Figure 70). The variant segregated with the disease in both families (Figure 70) and was located within a 14 Mb region of homozygosity. The variant was absent in publicly available population databases including gnomAD, Exome Sequencing Project (*ESP),* GME Variome and Iranome, as well as in over 14,000 in-house exomes (Table 23).

In family 3, a novel homozygous missense variant c.1070G>A (p.Gly357Asp) within a 10 Mb block of homozygosity was detected. The variant occurs at a highly conserved amino acid residue, CADD score of 27 and GERP score of 5.49 and segregated with the disease in the family. A novel segregated homozygous missense variant c.986C>G, p. (Pro329Arg), altering a highly conserved amino acid was identified for family 4 (GERP score of 5.37). Based on in silico predictions, both p.Gly357Asp and Pro329Arg, are damaging and deleterious variants.

In family 5, two *PIGS* variants were identified in patient 6. One variant is a paternally inherited in-frame insertion of eight amino acids in a non-repeat region, as c.1141_1164dup24 (p.Asp381_Val388dup) in exon 10. This variant has been previously reported in the literature in a male of Chinese descent with *PIGS*-related disorder and seen in gnomAD but only in a heterozygous carrier state in five individuals of East Asian ethnicity (frequency of 0.029%). The second variant is a maternally inherited nonsense variant c.734G>A (p.Trp245*) in exon 7. This variant

has been seen in gnomAD but only in the heterozygous carrier state in one individual of African ethnicity (frequency of 0.0115%).

Table 23. *PIGS* intragenic variants identified in our cohort.

	Patient ID	Probands F1-II:1 F2-II:1	Proband F3-II:1	Probands F4-II:1 F4-II:2	Proband F5-II:1	Proband F5-II:1
Variant annotation	GRCh37/hg19 position (DNA change)	chr17:26898067	chr17:26883855	chr17:26883939	chr17:26883200	chr17:26887152
	cDNA change (NM_033198.4)	c.174G>C	c.1070G>A	c.986C>G	c.1141_1164dup24	c.734G>A
	Protein change	p.(Gln58His)	p.(Gly357Asp)	p.(Pro329Arg)	p.(Asp381_Val388dup)	p.(Trp245*)
	Inheritance	Hom	Hom	Hom	Comp Het	Comp Het
	dbSNP ID	-	-	-	rs769890071	rs1249675321
	Variant seen in family	F1-F2	F3	F4	F5	F5
Allele frequencies (PM2)	gnomAD v3 (highest subpopulation)	-	-	-	0.029% (5/17248)	0.0115% (1/8732)
	gnomAD v2.1.1 (highest subpopulation)	-	-	-	-	0.011% (1/8714)
	Frequency in ensembl browser	-	-	-	-	-
	Iranome	-	-	-	-	-
	GME Variome	-	-	-	-	-
	Frequency in in-house database‡	-	-	-	-	-
	Frequency in GeneDx database	0.001% (1/139678), no homs	0.001% (1/130874), no homs	-	1/164434, no homs	0.001% (1/164434), no homs
***In silico* predictions (PP3)**	GERP	5.05	5.49	5.37	-	5.84
	CADD	25.2	27.4	28.9	22.3	38
	Polyphen-2	B (0.431)	PD (1)	PD (0.964)	-	-
	SIFT	T (0.09)	D (0)	D (0)	-	-
	Provean	N (-1.725)	D (-6.756)	D (-6.029)	D (-11.319)	D (-14.144)
	MutationTaster	DC (0.999)	DC (0.9999)	DC (0.9999)	-	DC (1)
ACMG	Overall classification	PP1, PM2, PM5, PS3	PM2	PP1, PM2	PM3, PM4, PP3, PP4	PVS1, PP4

Splicing analysis

Computational splice analysis performed of the c.174G>C variant on Alamut visual (owned by Dr. Barbara Vona, University of Tuebingen) predicted two possible effects that prompted RNA studies: 1) a decrease in the splice donor scores between 12.6%

and 43.3% and 2) the abolishment of an exon splice enhancer motif (Figure 73). *In vitro* RNA analysis of the wild-type c.174G disclosed a 508 bp amplicon including exons 2 and 3, as well as a 369 bp amplicon with only exon 3 (Figure 74). This likely reflects alternative splicing (NM_033198.4/ENST00000308360.8 and ENST00000395346.6). The amplicon with the homozygous c.174G>C variant revealed only exon 3 (369 bp), indicating a skipping of exon 2. This non-frame deletion of exon 2 (r.35_174del) would result in a premature stop codon (p.(Glu12Alafs*31)).

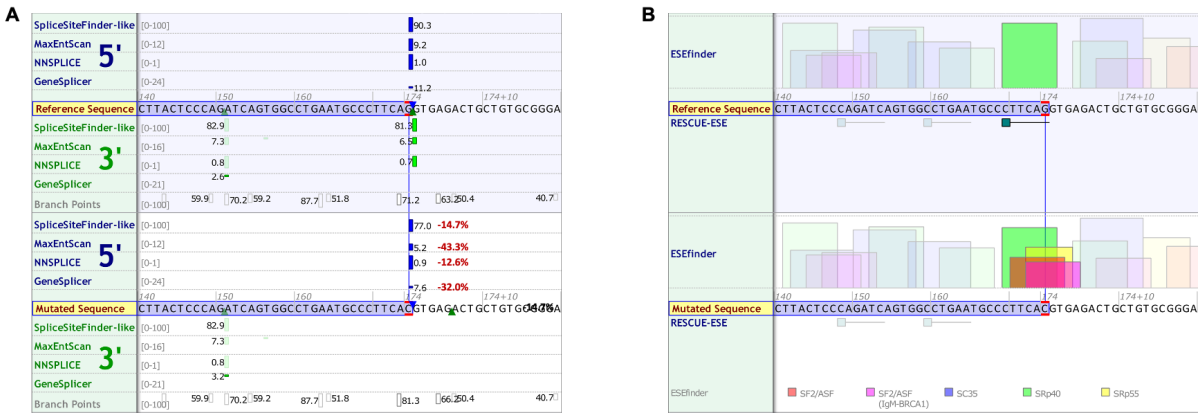


Figure 73. *In silico* splice prediction scores 277 of the *PIGS* c.174G>C variant. **A.** Four splice prediction tools using Alamut Visual compare the 5‘ and 3‘ splice site scores oft he wild-type (upper panel) and variant (bottom panel). The native splice donor site is maintained (blue bars), but all splice prediction scores decrease (lower panel, red). **B.** Analysis with ESEfinder and RESCUE-ESE tools reveal the splicing sequence landscape of the wild-type (upper panel) and variant (bottom panel) sequence. Changes in ESEfinder splicing factor predictions are shown above the sequence in coloured bars. An abolished ESE motif is depicted by a loss of a green box below the sequence. The c.174 nucleotide is outlined in red bars for both (A) and (B).

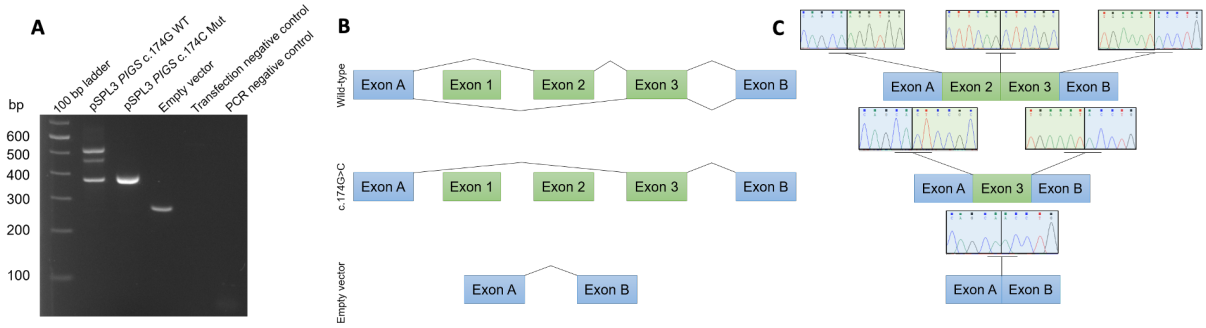


Figure 74. RNA analysis of the *PIGS* c.174G>C variant. A. Gel electrophoresis of the RT-PCR wild-type (WT), mutant (Mut) c.174G>C variant, and empty pSPL3 vector amplicons. Transfection negative and PCR negative controls performed as expected. **B.** Overview of the splicing of the wild-type (upper panel), patient (middle panel) and empty vector (lower panel) splicing. Note that although exon 1 was cloned into the vector, it was not spliced to exon A and exon 2, as it lacks a splice acceptor site. **C.** Visualization of the exon junction sequence of the wild-type (upper panel), patient (middle panel) and empty vector (lower panel) (data generated by Barbara Vona).

Flow Cytometric Analysis

Individual 6 had significantly lower level of CD16 and FLAER in granulocytes, the most sensitive markers of Inherited GPI Deficiency (Figure 75). There was no decrease for CD55 and CD59 in granulocytes but these are less sensitive markers of Inherited GPI Deficiency. There was also low CD14 in monocytes and low FLAER in lymphocytes (Figure 76). This is consistent with the previously reported case series that showed biallelic variants in *PIGS* lead to reduced amounts of GPI-AP at the cell surface.[15-16]

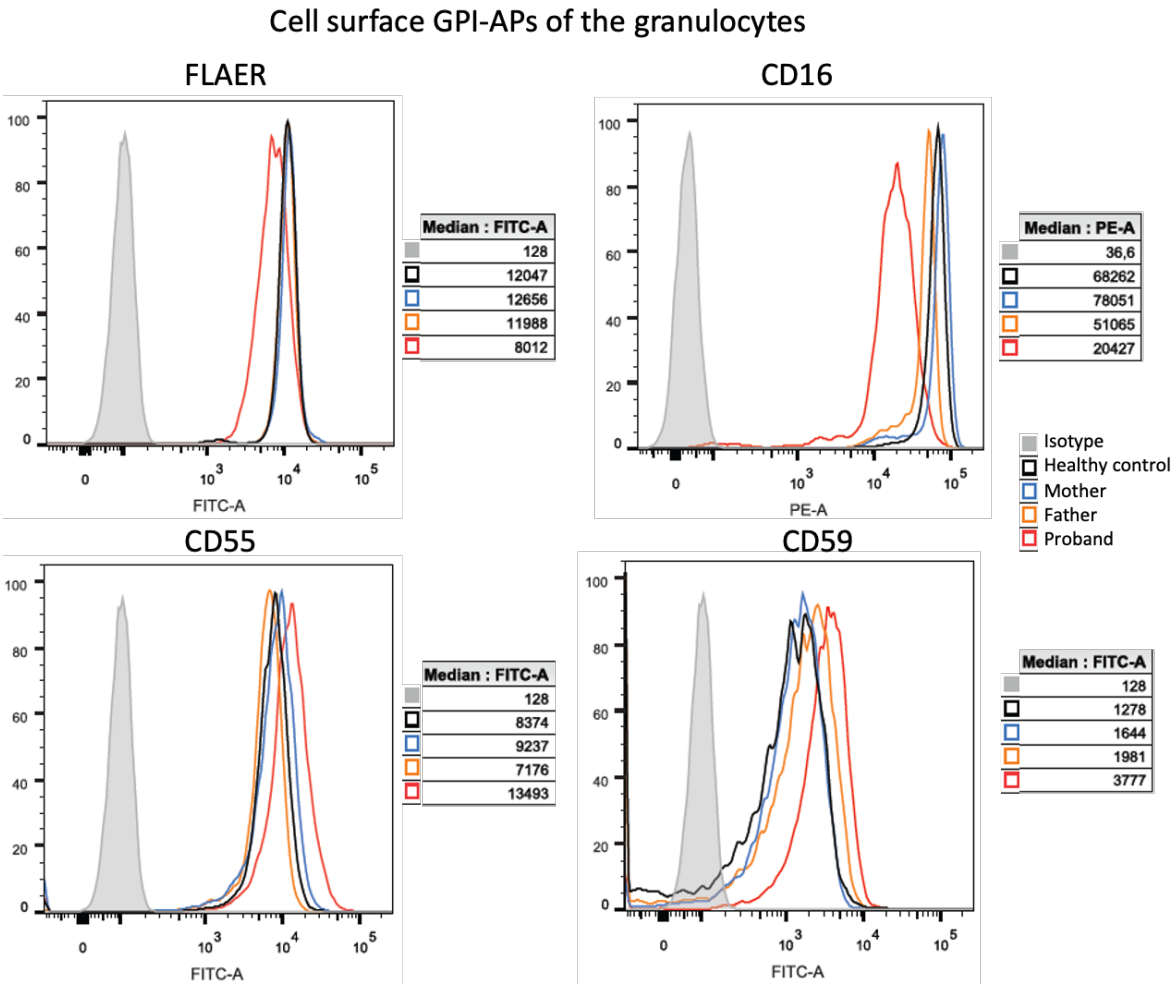


Figure 75. Impact of the *PIGS* variants on individual granulocyte cell-surface GPI-APs. Red blood cells were lysed in BD FACS lysing solution (BD Biosciences) from fresh blood of patient 6 from family 5 and control cells, and then the samples were stained with GPI-AP markers (FLAER, CD16-FITC, CD55-FITC, and CD59-FITC) for 20 min on ice. The nonspecific binding was washed before analysis by the BD FACSCanto II system (data generated by Philippe Campeau).

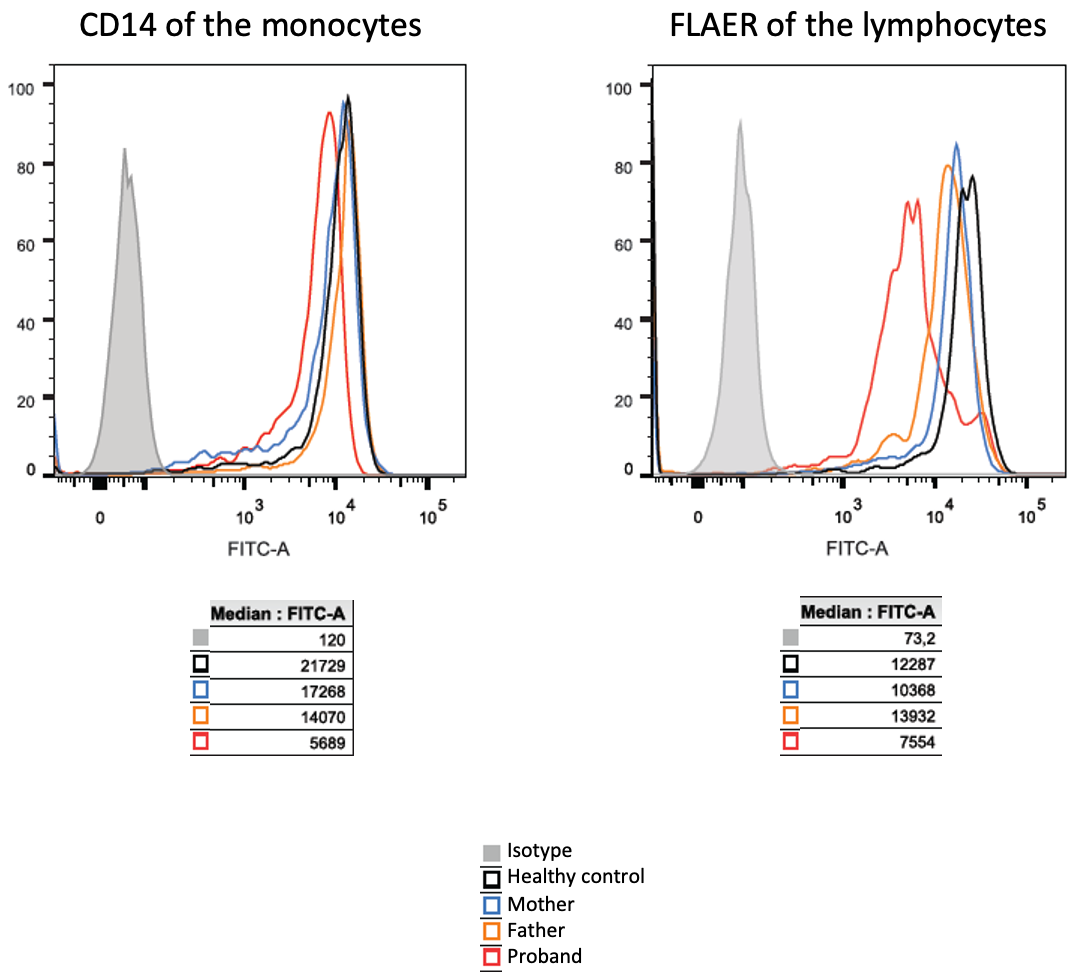


Figure 76. Impact of the *PIGS* variants on individual lymphocyte and monocyte cell-surface GPI-Aps. Flow-cytometry analysis of lymphocytes and monocytes from the same experiments described in Figure 2. Shown is a representative analysis of the amount of cell-surface CD14 on monocytes and FLAER on lymphocytes from triplicate experiments (data generated by Philippe Campeau).

5.2.3 Discussion

Our study expands the phenotypic and genotypic spectrum of *PIGS* variants, which have been consistently associated with an epileptic-dyskinetic encephalopathy with features of early infantile epileptic encephalopathy, profound disability, and premature death. We expand the known population of 7 individuals with biallelic

pathogenic variants in *PIGS* by adding 5 new patients. We report 5 variants, 4 of which were previously unreported; namely, one truncating, one in-frame duplication and three missense variants. According to gnomAD constraint scores, *PIGS* is intolerant towards missense variants (Z = 0.7801) and biallelic loss of function variants (pLI=0, pREC=0.9943).

All of our patients had infantile-onset epilepsy with hypotonia, microcephaly, global developmental delay, and no speech. We discovered additional neuroimaging findings in our patients, consistent with previously published studies that described cerebellar and diffuse cortical atrophy in individuals with *PIGS* variants.[274; 275] In this cohort, vision impairment, hearing loss, and behavioral abnormalities were common. Taken together, biallelic gene-disruptive variants in *PIGS* are associated with an early infantile encephalopathy difficult to treat with standard AEDs co-occurring with global developmental delay, severe hypotonia, and visual and hearing impairment.

Two of our patients were treated with pyridoxine and folinic acid with some improvement on neurological status or epilepsy. Although anectodal, previous studies have also shown the effects of pyridoxine in other IGDs.[276] Pyridoxine has been shown effective in various GPI-deficiency disorders, though there is no clear evidence describing the effect of folinic acid. Ketogenic diet has been previously reported to improve outcome in cases of *PIGA* deficiency and may be a consideration for individuals with *PIGS*-related disorder.[277]

This study further reviewed the clinical features of *PIGS*-related disorders. Distinguishing *PIGS*-related disorders from other paediatric neurological syndromes

will rely on the combination of physical exam findings and identifying the phenotypic features resulting from disruption of *PIGS*, including epilepsy, developmental delay, hypotonia, and coarse facies. Our data further stress the importance of including the screening of the *PIGS* gene in the case of infantile epilepsy and may serve for improved interpretation of new *PIGS* variants with the help of biochemical findings and flow-cytometry analysis. Ultimately, *PIGS* is not found on most clinically utilized epilepsy gene panels suggesting that it may be a more frequent cause of early infantile epilepsy than reported. As epilepsy genetic testing expands and WES becomes more available, more patients will likely be reported.

5.3 Human variants in *SLITRK3* implicated in GABAergic synapse development

5.3.1 Introduction

Synapse development is a multi-step process coordinated by various molecules acting in a highly spatially and temporally controlled manner. Synaptic cell adhesion molecules (CAMs) are an example of such molecules essential for the establishment and maturation of synaptic connections. Multiple cell-adhesion molecules contribute to synapse formation by mediating trans-synaptic interactions with presynaptic signalling molecules.

GABAergic synapses provide inhibition in the mammalian brain, providing an important level of inhibitory balance to excitatory glutamatergic drive therefore controlling neuronal excitability, synapse formation and plasticity. As GABAergic inhibition is important in almost every aspect of brain physiology and the dysregulation of GABAergic synapse development has been implicated in many neurological, neurodevelopmental and psychiatric disorders [278; 279], it is critical to understand the molecular determinants of GABAergic synapse formation. Prevention and treatment of these diseases will depend partly on restoration of GABAergic function and/or inhibitory/excitatory balance.

However, GABAergic transmission can no longer be considered solely inhibitory, since Cl^- efflux, which itself is dependent on the expression and activity of cation chloride transporters, is depolarizing. These findings have profoundly influenced the consideration given to GABAergic transmission in adult and developing CNS. It is

now widely accepted that GABAergic synapses are functionally heterogeneous, precisely determine the function of neuronal networks, and control neuronal development and differentiation during ontogeny and in adult neurogenesis.[280]

Recent studies have identified a growing number of synaptic CAMs such as slit and NTRK-like family proteins (Slitrks) that bind to presynaptic protein tyrosine phosphatase (Figure 77). Members of the Slitrk family, such as ST3, are integral membrane proteins with two N-terminal leucine-rich repeat (LRR) domains. They also have C-terminal regions that share homology with neurotrophin receptors. They are expressed predominantly in neural tissues and have neurite-modulating activity.[281] Even though such molecules are involved in various stages of synapse development, their function diversity still remains largely unclear.

Single-passing transmembrane protein, Slitrk3 (Slit and Trk-like family member 3, ST3), is a synaptic cell adhesion molecule highly expressed at inhibitory synapses. Little was known until recently about the physiological function of ST3 intracellular, carboxyl (C)-terminal region. A recent study by Li *et al*. (2017) has added a major novel element to our understanding of the mechanisms of GABAergic synapse formation by revealing a direct extracellular protein-protein interaction between ST3 and another postsynaptic cell-adhesion molecule, Neuroligin 2 (NL2). They have shown that ST3, through its extracellular domain, selectively regulates inhibitory synapse development via the trans-synaptic interaction with presynaptic cell adhesion molecules, neurexins or receptor protein tyrosine phosphatase δ (PTPδ) and the *cis*-interaction with postsynaptic NL2. In this way, they exert differential,

sometimes cooperative effects on GABAergic synapse formation depending on the developmental stage of the system (Figure 77).

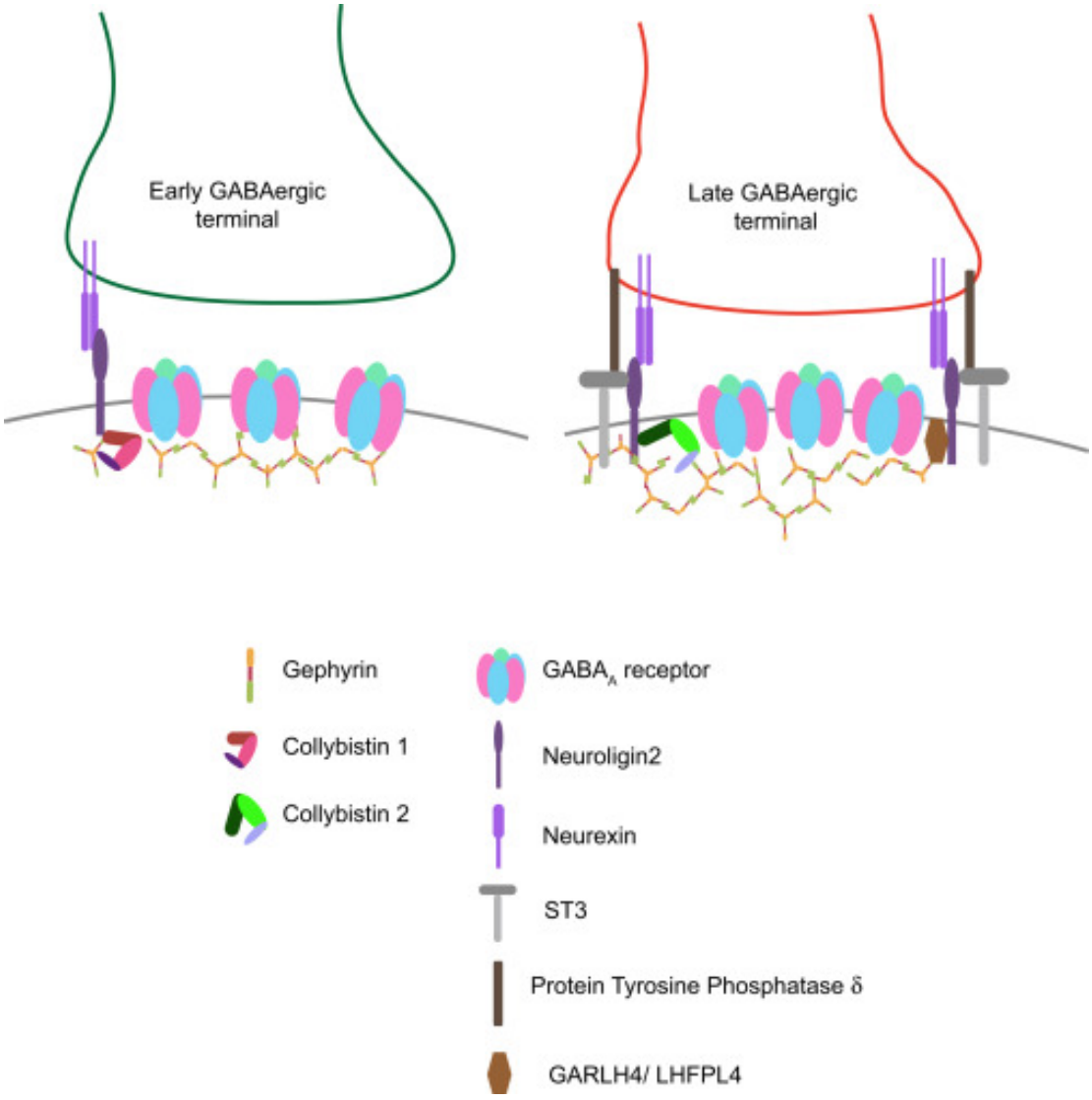


Figure 77. Role of NL2 and ST3 during Early and Late GABAergic. CB1 interaction with NL2 in developing neurons is sufficient to drive GABAergic synapse maturation, whereas in mature neurons CB2 interacts with the NL2/ST3 complex to facilitate GABAergic synapse maturation (Fritschy *et al.,* 2017).

In the brain, many types of interneurons make functionally diverse inhibitory synapses onto principal neurons. Li *et al*. (2019) have implicated Neuroligin2 (NL2) and SLITRK3 as two synaptogenic molecules critical for GABAergic synapse

development in the hippocampus. They reported that in heterologous cells, ST3 C-terminus is not required for ST3 homo-dimerization and trafficking to the cell surface. In contrast, in hippocampal neurons, the ST3 C-terminus is critical for GABAergic synapse development. Indeed, overexpression of an ST3 C-terminus (Y969A) mutant markedly reduced the gephyrin puncta density and GABAergic transmission in hippocampal neurons. In addition, single-cell genetic deletion of ST3 strongly impaired GABAergic transmission. Importantly, wild-type (WT) ST3, but not the ST3 Y969A mutant, could fully rescue GABAergic transmission deficits in neurons lacking endogenous ST3, confirming a critical role of Y969 in the regulation of inhibitory synapses and showing the importance of a single critical residue in ST3 C-terminus for GABAergic synapse development and function.[282; 283]

Li *et al*. (2017) report that in DIV8 neurons the CB-binding domain of NL2 is required for GABAergic synaptogenesis, while in DIV18 neurons NL2 interaction with ST3 is essential for GABAergic synaptogenesis and synapse development. This temporal dependence on NL2 and ST3 coincides with the temporal dependence of collybistin (CB) splice isoforms CB1 and CB2, which in turn are regulated by alternative splicing and distinct subcellular localisation. CB1 is expressed early during postnatal development and influences GABAergic synapse formation *in vitro* along the proximal-distal axis of dendrites[284]. Separately, CB2 expression is not developmentally regulated, and it forms no proximal-distal gradient of GABAergic synaptogenesis.

Here we report on biallelic and monoallelic mutations in *SLITRK3* in 3 families presenting with epileptic encephalopathy associated with neurodevelopmental

findings. The primary cultures of hippocampal neurons carrying ST3 variants were prepared in collaboration with Dr. Wei Lu's lab (NIH, Bethesda) through a combination of cloning, transfection of the mouse hippocampal neurons, electrophysiology (and immunocytochemistry in the future) to try to characterize the neuronal impact of these variants.

5.3.2 Results

Clinical findings

Family 1 consists of two 7-year-old (F1:II-1) and 6-year-old (F1:II-2) boys and a 5-year-old girl (F1:II-3) from a consanguineous family of Pakistani descent (parents are first cousins) (Figure 78). Family history was unremarkable, except for one prior spontaneous miscarriage. The pedigree suggested an autosomal recessive inheritance. All presented with generalised tonic-clonic seizures since a young age (6 months for F1:II-1, 24 months for F1:II-2, and 18 months for F1:II-3). They all developed neurodevelopmental delay which includes speech delay or absence and intellectual disability. F1:II-2 and F1:II-3 have short stature; 94cm (<0.4th centile) 84cm (<0.4th centile), respectively. F1:II-3 also has microcephaly at 44cm (<0.4th centile) EEG analyses showed left parietal focal epileptogenic activity in F1:II-1 and diffuse encephalopathy in his two other siblings. Ophthalmological examination in F1:II-1 showed RPE abnormalities. The clinical data was provided by Dr. Faisal Zafar.

Family 2 consists of a 5-year-old boy of (F2:II-1) of American origin (Figure 78). The pregnancy of the child was complicated by a IUGR but was born full term. He first presented with febrile seizures at 3 years of age accompanied by generalized fever

(102°F) with stiffness, focal shaking and limping, and which lasted less than one minute. Six months later he experienced generalised tonic-clonic seizures with post-ictal symptoms. Both episodes were postictal. He has experienced both febrile and non-febrile seizures since. Overnight video EEG done at the age of 2 years showed background slowing, occasionally max left temporal, multifocal IEDs, max right central, often admixed with sleep architecture. Brain imaging done at the age of 2 years showed multiple subcortical white matter T2 hyperintensities without enhancement or mass effect. These lesions were nonspecific and the differential diagnosis was long including gliosis, demyelination, dysmyelination, Lyme, vasculitis and chronic ischemia. There was optic nerve head flattening suspicious for papilledema. No developmental concerns and a motor skill assessment revealed no atrophy or fasciculation but a mild low tone. Ophthalmology evaluated him and dilated exam showed normal healthy optic nerves without any edema. The patient was alert and interactive, he made good eye contact, and followed commands. Speech was fluent and language was appropriate for his age. Recent and remote memory are grossly intact. The clinical data was provided by Dr. Shenela Lakhani.

Family 3 consists of a 16-year-old girl (F3:II-1) of African American origin (Figure 78). The pregnancy of the child was complicated by IUGR but she was born full term. Parents were concerned about motor delays very early on but became very concerned around the age of 9-12 months. She was diagnosed with epilepsy, severe global developmental delay and intellectual disability. She rolled over at 6 months but did not meet any other motor milestones until after 2 years old and she did not walk until 6 years old. She currently remains nonverbal and does not use signs; she communicates via vocalizations and walking to what she wants. She started having

staring spells at around 9-12 months old. These progressed to generalised tonic-clonic seizures lasting up to 2 minutes with post-ictal symptoms. Her seizures have been controlled with Leviteracetam and Clonazepam. Her last seizure was 18 months ago. F3:II-1 has short stature (<0.4th centile), microcephaly at 47cm (<0.4th centile) and she has dysmorphic features (larger ears, prominent nose, larger mouth, short philtrum, everted upper lip, camptodactyly in fingers, flat feet). She has also been diagnosed with cortical vision impairment and nystagmus. The mother also reports significant behavioural issues including breath holding, difficulty staying asleep, teeth grinding, hand rubbing, laughing fits, and crying spells. Brain imaging done at the age of 11 months showed mild ventricular enlargement of prominent sulci suggestive of low brain volumes, thinned corpus callosum, right mastoid air cell fluid and an irregular border to the ventricular system, raising the possibility of periventricular leucomalacia. Other clinical features included hypotonia and spasticity. Details clinical details of the patients identified in our cohort can be seen in the Appendix (Supplementary Table 10). The clinical data was provided by Dr. Li Hong.

Genetic findings

In total, 83,572,847 (II-1) and 81,527,162 (II-2) unique reads were generated. After applying the previously mentioned filtering criteria (see Methods for Chapter 5.3), no plausible shared compound heterozygous variants were identified by WES; there was however a gene carrying a rare (likely) damaging variant, according to guidelines for variants interpretation[71], which were homozygous in all three probands in family 1 (Figure 78). A homozygous frameshift deletion in *SLITRK3* (NM_014926:c.G1816T:p.Glu606*) emerged as the most likely explanation for the

disease pathogenesis; this is also supported by a more severe impact of the mutation on protein function (truncating vs. missense) and an existing functional report previously linking cell adhesion molecules (CAM) such as SLITRK3 to synaptogenesis and GABAergic synapse development[282]. Segregation analysis performed by traditional Sanger sequencing confirmed the homozygous variant in the three affected siblings and heterozygous in both their parents. The identified *SLITRK3* homozygous variant was submitted to the Leiden Open Variation Database (www.lovd.nl/; variant ID #0000693802). Through clinical whole exome sequencing carried out at Genedx, a *de-novo SLITRK3* variant (c.1696T>C, p.Cys566Arg) was identified for family 2 and a *de-novo* mosaic frameshift mutation (c.660_684del25, p.Gly221TrpfsX15) in family 3.

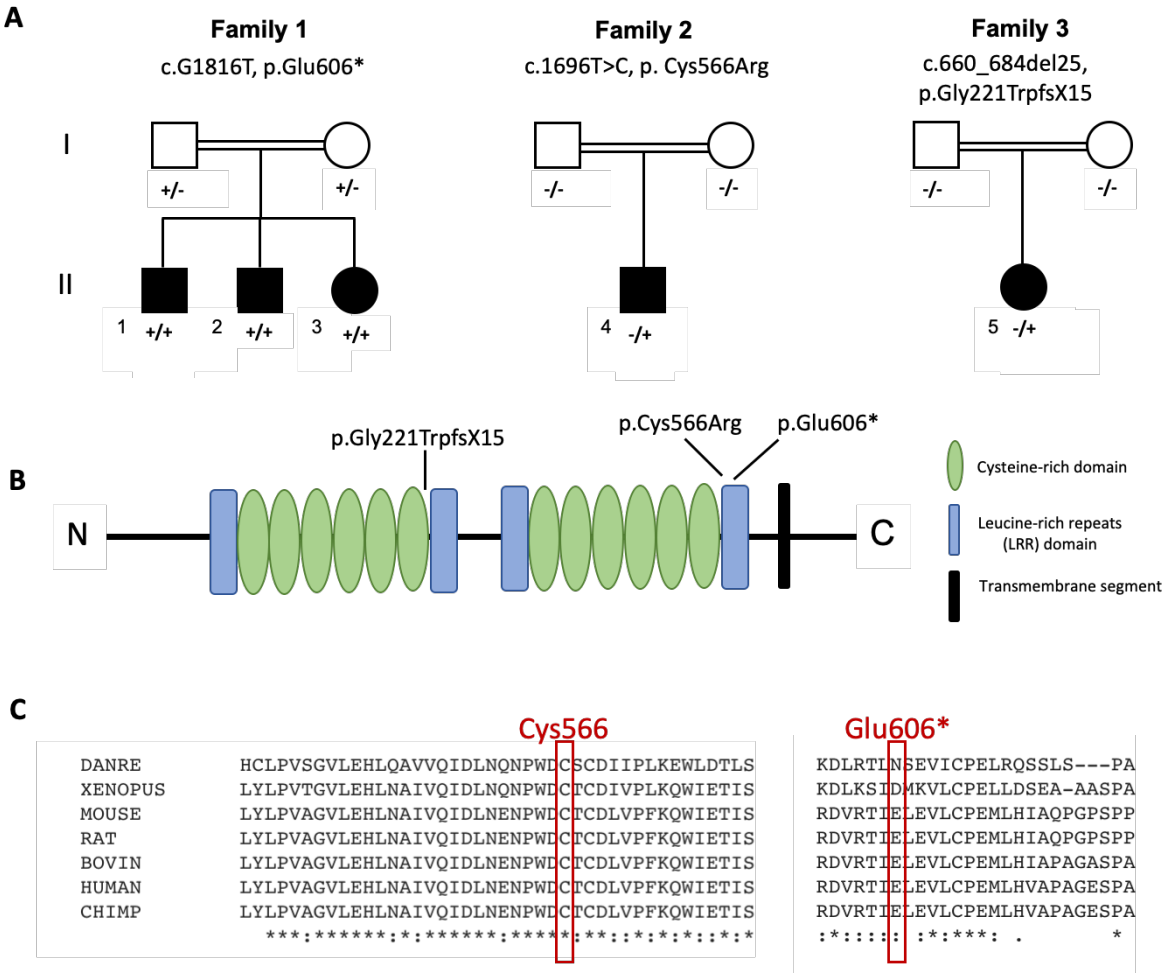


Figure 78. Pedigrees and genetic findings of the families carrying biallelic *SLITRK3* variants. A. Pedigrees and segregation results (+ represents the presence of the variant) of the three families **B.** A schematic representation of the ST3 protein showing the position of all variants identified. **C.** Inter-species alignment performed with Clustal Omega shows the complete conservation down to invertebrates of the amino acid residues affected by the substitutions.

ST3 protein expression

In HEK293T cells, using two antibodies (anti-Myc and anti-Slitrk3) and also N-terminal Myc-Slitrk3 as a positive control, Western blot experiments showed that only candidate p.Glu606* strongly decreased the expression of co-transfected Myc-ST3 (Figure 79).

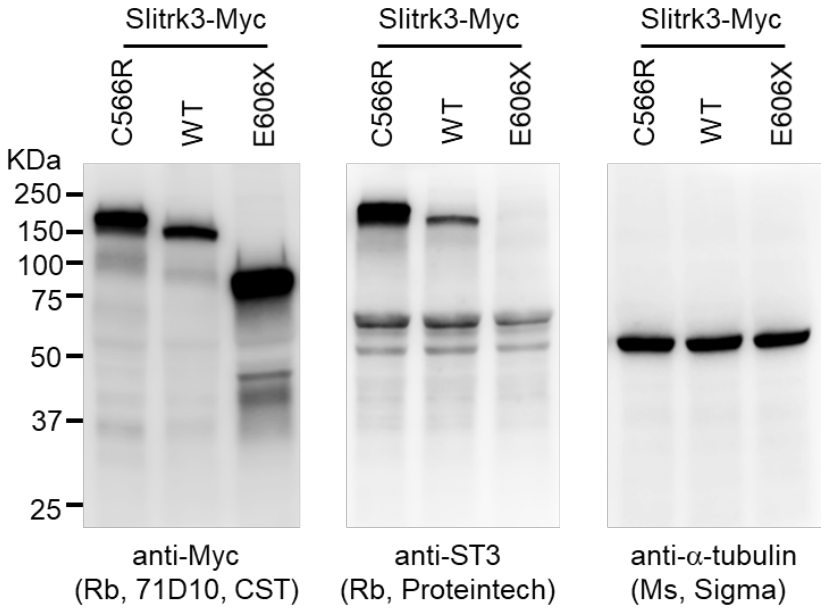


Figure 79. Western blot analysis showed that the frameshift variant p.Glu606* can significantly reduce the expression levels of ST3. C566R denotes the plasmid carrying the p.Cys566Arg variant, E660X denotes the plasmid carrying the p.Glu606* variant, WT denotes wilt-type. ST3 is alternatively used for SLITRK3 for

short in the protein, e-phys, ICH studies. Expected size band for ST3 is 120-125kDA. α-tubulin was used as an internal control. N = 3 independent repeats.

Whole cell recording of human ST3 mutants

We performed the whole-cell recordings to measure miniature inhibitory postsynaptic currents (mIPSCs) in cultured hippocampal neurons (DIV14) overexpressing human WT-ST3, ST3-C566R and ST3-E606X mutants (also simultaneously expressing GFP). In total, around 10 cells for each plasmid group were analysed from six different batches. In neurons overexpressing WT-ST3, the frequency, but not amplitude, of mIPSCs was significantly increased (Figure 80) compared with the GFP-negative cells. Based on previous studies, overexpression of ST3-FL should increase the mIPSC frequency, and that overexpression of ST3-FL increased the expression level of Gephyrin.[282; 285]

In contrast, in neurons overexpressing ST3-C566R, GABAergic transmission had no change compared with the control cells (Figure 80) indicating an essential role of ST3-C566R in the regulation of inhibitory transmission. In addition, the mIPSC frequency and amplitude had no change in ST3-E606X mutant compared with the control cells. The E606X is a truncation mutant without the transmembrane domain and all C-terminus, therefore the phenotype in ST3-E606X could be due to a loss of WT-ST3.

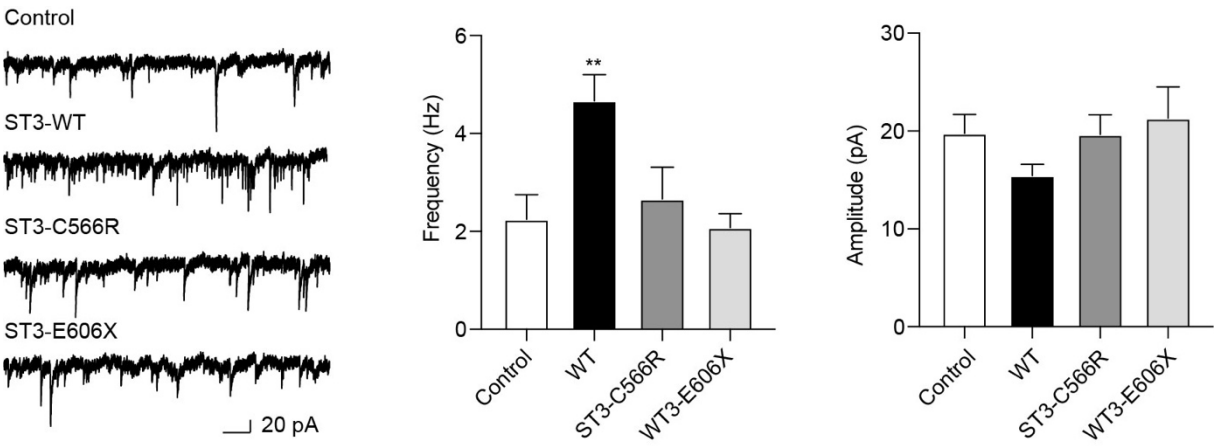


Figure 80. Human mutants in ST3 is essential for GABAergic synaptic transmission. mIPSC recording showed that overexpression of human WT-ST3 significantly increased mIPSC frequency, whereas overexpression of the ST3-C566R mutant had no effect on the frequency of mIPSCs in cultured hippocampal neurons. Insets showed the mean ± SEM of mIPSC frequency and amplitude [Frequency (Hz): Control, 2.25 ± 0.50, n = 10; WT-ST3, 4.67 ± 0.54, n = 9; ST3-C566R, 2.67 ± 0.65, n=11; ST3-E606X, 2.08 ± 0.28, n=10. One-way ANOVA test, ** p = 0.0058; Mann-Whitney test, Control vs ST3-WT: ** p = 0.004; Control vs ST3-C566R: p = 0.918; Control vs ST3-E606X: p = 0.579. Amplitude (pA): Control, 19.770 ± 1.91, n = 10; ST3-WT, 15.44 ± 1.14, n = 9; ST3-C566R, 19.65 ± 2.02, n=11; ST3-E606X, 21.30 ± 3.22, n=10. One-way ANOVA test, p = 0.33; Mann-Whitney test, Control vs ST3-WT: p = 0.08; Control vs ST3-C566R: p = 0.76; Control vs ST3-E606X: p = 0.91. N = 5-6] Scale bar, 20 pA and 1 s (data generated by Wenyan Han).

5.3.3 Discussion

The expression of human SLITRK genes has been detected mainly in the brain, but the expression profile of each SLITRK is unique. SLITRK3 is expressed in the

occipital lobe of the cerebral cortex of the brain. Interestingly, SLITRK3 expression is enhanced in tissue from several different types of brain tumors such as astrocytomas, oligodendrogliomas, glioblastomas, gangliogliomas and primitive neuroectodermal tumors.[281]

Genetic mutations in other SLITRK genes like *SLITRK1, SLITRK2* and *SLITRK4* have been associated with a multitude of neuropsychiatric disorders such as schizophrenia and autism spectrum disorders. However, very little is known about the neuronal and synaptic consequences of these mutations.[286] Here, we report on biallelic and monoallelic mutations in *SLITRK3* in 3 families presenting with epileptic encephalopathy associated with a broad neurological involvement characterized by microcephaly, intellectual disability, seizures, and global developmental delay. *SLITRK3* encodes a transmembrane protein that is involved in controlling neurite outgrowth and has an important role in brain function and neurological diseases.

We aimed to characterise the neuronal phenotype of mouse-derived hippocampal neurons by transfecting them with mutant ST3 plasmids and conducted electrophysiological and cell biological analyses of GABAergic synapses. Preliminary results showed that mutants ST3-C566R and ST3-E606X exhibit impaired development of peripheral and central nervous system through a combination of loss of WT-ST3 protein and a loss of function effect. Among the two mutants, p.Glu606* and p.Cys566Arg, the latter is especially interesting biologically. It appears that the frameshift mutation p.Glu606* mutation is located in the transmembrane domain, leading to a truncated ST3 protein without a transmembrane or C-terminus tail. We expect the extracellular residue stretch to not be able to anchor on the surface. Thus,

while WT ST3 is a transmembrane protein, this mutant will be an intracellular, cytosolic protein, or a secreted protein if the signal peptide is kept in the construct. Thus, the phenotype in p.Glu606* could be due to a combination of deficits, loss of WT SLITRK3 and a possible mi-localization. Upon investigating this mutant further, we found that an overexpression does not have a real physiological effect on the synaptic transmission but further immunostaining studies will shed more light on the role of this variant in GABAergic synapse development in the transfected neurons.

In contrast, p. Cys566Arg is more interesting in terms of understanding the role of ST3 in GABAergic synapses. Overexpression of the ST3-C566R mutant had no effect on the frequency of mIPSCs in cultured hippocampal neurons and no effect on GABAergic transmission. Since the p.Cys566Arg variant will probably produce a full-length ST3 mutant, we speculate that it might cause a loss of function effect and act as a dominant-negative mutation. This indicates a critical function of the domain where Cys566 resides and an essential role of ST3-C566R in the regulation of inhibitory transmission. We plan to investigate further by immunostaining studies.

Taken together, our results support a conserved role of this transmembrane protein in neurological function and an essential role of ST3-C566R in the regulation of inhibitory transmission. Given the previous study by Li *et al.* (2017) which already showed that ST3 C-terminus is critical for GABAergic synapse development, we are interested to further observe gephyrin puncta density, surface staining and GABAergic transmission in hippocampal neurons.

Chapter 6. General conclusions

In this PhD thesis, I used a combination of high-throughput genetics and molecular biology to define new approaches to advance understanding and disease characterisation in rare neurological disorders. Using these strategies, my work has helped to discover three new disease-causing genes (*NFASC, NARS1, SLITRK3*), two of which have since been independently validated by other laboratories[2; 4; 287-289], and one of them (*NFASC*) has been catalogued as a new disease by the Online Mendelian Inheritance in Man (OMIM # 618356).

In all the gene discoveries, I demonstrate the importance of independent validation and the vital role of international collaboration in studying rare diseases. The future of genetics lies in high-throughput sequencing, and clinical geneticists are rapidly embracing whole exome and whole genome sequencing as diagnostic tools. It is anticipated that in the next few years, the rate of gene discovery will continue to accelerate, with an exponential explosion of novel genetic findings.

New disease-gene discoveries

In Chapter 3, I have shown that rare diseases genes can be associated with previously known phenotypes which we did not think could be hereditary before. The first new disease presented here was *NFASC*- associated neurodevelopmental impairment and demyelination (OMIM # 618356). We have classified it as a hereditary nodo-paranodopathy, a newly recognised autosomal recessive neurodevelopmental disorder clinically characterised by progressive infantile demyelinating neuropathy, hypotonia and developmental delay. This discovery and functional validation were possible due to the integration of basic research in the field

of high-throughput genome analysis with biochemistry and cell biology via the establishment of strong collaborations. In order to identify the disease-causing gene and to establish the disease mechanism, we initially performed whole exome sequencing in the affected case and prioritized the filtering strategy for novel and very rare variants with high impact on the protein. We then used homozygosity mapping, further filtered the variants in the shared region and identified one in the NFASC gene.

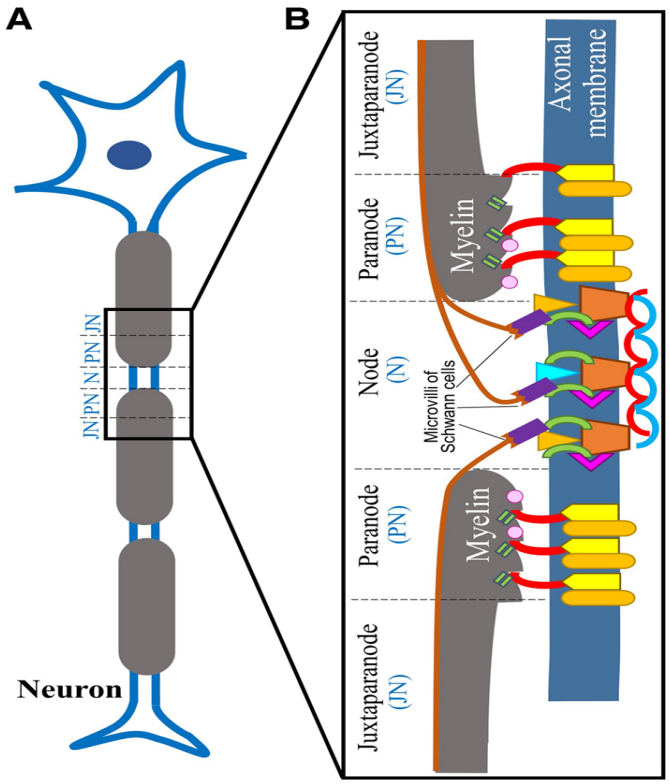
Using Gene Matcher and connections with other researchers, we identified the biggest cohort to date of biallelic variants in *NFASC* in 10 individuals from six independent families with global developmental delay and muscular weakness with variable features of central hypomyelination and chronic demyelinating neuropathy. In comparison to previous case reports, [2; 4] the mutations described by us are in different neurofascin isoforms and in different combinations, such as NFASC155 and NFASC186 together, NFASC186 alone, or NFASC155 alone. Individuals with variants affecting the NFASC155 isoform have disease onset at birth or within 6 months of life, whereas individuals with variants affecting only the NFASC186 isoform have later disease onset.

To study the impact of *NFASC* variants at the cellular level, we examined the surface expression and the protein levels of each variant. We found reduced surface expression and NFASC levels only with the loss-of-function variant in NFASC155 (p.Pro939Ter). Only those variants affecting the fibronectin type III domain were shown to reduce protein levels significantly. We also revealed a disrupted interaction between neurofascin and paranodal proteins CNTN1/CASPR1 in N2A cells. This

impaired interaction between paranodal proteins is another important piece of evidence for the causal role of the NFASC variants, as a similar phenotype (congenital hypomyelinating neuropathy, OMIM 618186) was shown in individuals with disrupted CASPR1 function caused by variants in *CNTNAP1*.[133]

Immunofluorescence staining showed a severe loss of myelinated fibres around hair follicles and a complete sparing of unmyelinated fibres in the epidermis and dermis. We concluded that the abundant expression of NFASC in newly formed oligodendrocytes could be the factor underlying the neurodevelopmental problems of individuals with biallelic *NFASC* variants.

We have provided a good example of how similar diseases can emerge from different aetiologies. Hereditary nodo-paranodopathies are demyelinating polyneuropathies with a monogenic aetiology (Figure 81).



C

Symbol	Protein	Gene	Disease (OMIM entry)
Paranodopathies			
	Neurofascin Isoform 155	*NFASC*	Neurodevelopmental disorder with central and peripheral motor dysfunction (618356)
	Contactin 1	*CNTN1*	?Myopathy, congenital, Compton-North (612540)
	Contactin associated protein 1	*CNTNAP1*	Hypomyelinating neuropathy, congenital, 3 (618186) Lethal congenital contracture syndrome 7 (616286)
	Connexin 32	*GJB1*	Charcot-Marie-Tooth neuropathy, X-linked dominant 1 (302800)
Nodopathies			
	Neurofascin Isoform 186	*NFASC*	Neurodevelopmental disorder with central and peripheral motor dysfunction (618356)
	Gliomedin	*GLDN*	Lethal congenital contracture syndrome 11 (617194)
	Beta-IV-Spectrin	*SPTBN4*	Neurodevelopmental disorder with hypotonia, neuropathy, and deafness (617519)
	Alpha-II-Spectrin	*SPTAN1*	Epileptic encephalopathy, early infantile, 5 (613477)
	Ankyrin-G	*ANK3*	?Mental retardation, autosomal recessive, 37 (615493)
	Myelin-associated Glycoprotein	*MAG*	Spastic paraplegia 75, autosomal recessive (616680)
	Sodium Voltage-Gated Channel, Alpha Subunit 1	*SCN1A*	Epilepsy, generalized, with febrile seizures plus, type 2 (604403); Dravet syndrome (607208); Febrile seizures, familial, 3A (604403); Migraine familial hemiplegic, 3 (609634)
	Sodium Voltage-Gated Channel, Beta Subunit 1	*SCN1B*	Epilepsy, generalized, with febrile seizures plus, type 1 (604233); Epileptic encephalopathy, early infantile, 52 (617233); Atrial fibrillation, familial, 13 (6157377), Brugada syndrome 5 (612838)
	Sodium Voltage-Gated Channel, Alpha Subunit 8	*SCN8A*	Cognitive impairment with or without cerebellar ataxia (614306); Epileptic encephalopathy, early infantile, 13 (614558); Seizures, benign familial infantile, 5 (617080)
	Potassium Channel, Voltage-Gated, KQT-like Subfamily member 2	*KCNQ2*	Epileptic encephalopathy, early infantile, 7 (613720); Myokymia (121200); Seizures, benign neonatal, 1 (121200)
	Potassium Channel, Voltage-Gated, KQT-like Subfamily member 3	*KCNQ3*	Seizures, benign neonatal, 2 (121201)
Nodal/Paranodal proteins not related to any disease			
	Neuronal Cell Adhesion Molecule	*NRCAM*	n/a

Figure 81. Nodal, paranodal and juxtaparanodal proteins of the peripheral myelinated axon, and overview of the hereditary nodo-paranodopathies. (A) Gross appearance and (B) detailed depiction of nodal/paranodal structure in longitudinal cross section. For simplicity, proteins have been excluded if they are expressed solely in the CNS, are located in the Schwann cell basal lamina, or are without known clinical relevance. (C) List of the genes encoding each of the proteins in B, plus any associated hereditary disorders. Neuronal cell adhesion molecule (NRCAM) is included for completeness but has not been associated with any disease (from Mert Karakaya and Brunhilde Wirth, 2019)

In contrast to the autoimmune polyneuropathies, hereditary nodo-paranodopathies often present with cortical findings (i.e. seizures, intellectual disability) due to high levels of protein expression in oligodendrocytes. Disease onset also differs between the inherited and autoimmune aetiologies, as the effects of biallelic mutations in NSAFC manifest at very young ages in contrast to the autoimmune aetiologies where disease onset is mostly in adulthood. As well as the later disease onset and symptoms restricted to the peripheral nervous system, autoimmune nodo-paranodopathies have another advantage in terms of therapy, with a partial to good response to rituximab or plasmapheresis. [290]

We have thus presented the largest clinical cohort to date of patients with *NFASC* mutations, and provided functional evidence implicating the major protein isoforms, which were also shown to be the main targets for the autoantibodies in CIDP pathogenesis. *NFASC*-associated neurodevelopmental impairment and demyelination should be considered a new member of the 'hereditary nodo-paranodopathies'.

In chapter 4, I have shown that rare diseases genes such as *NARS1* and *TARS1* that can encode for aminoacyl-tRNA synthetases in both mitochondria and the cytosol can be linked to a variety of recessive and dominant tissue-specific disorders and associated with different spectrum of phenotypes. AsnRS2, a mitochondrial ARS protein coded by *NARS2*, has recently been linked with an overlapping phenotype consisting of multisystem mitochondrial disorder (MID). Intellectual disability, epilepsy in childhood, hearing loss, and myopathy have also been seen in *NARS1* individuals described in this chapter.

A similar approach as in Chapter 3 was used here to identify the disease-causing gene as *NARS1* in the index proband recruited in Prof. Houlden's clinic. Using collaborative efforts and Gene Matcher we managed to collect thirty-two individuals with a neurodevelopmental phenotype with either *de-novo* heterozygous or biallelic mutations in the cytoplasmic *NARS1*. Studying these variants over the course of the last two years, has shown that their clustering and associated phenotypes at the N- and C-terminus are suggestive that these regions are functionally important and disrupt the protein homodimer and ATP-binding/catalytic domain in *NARS1.*

The N-terminal extension domain of *NARS1* has additional non-translational functions, which are common of such ARS proteins. By examining protein expression and synthesis, and also the enzyme's aminoacetylation activity, it is clear that such domains as the UNE-N, have evolved and don't seem necessary for ARS activity.

By using a combination of cell biology using either individual cell lines or neural progenitor cells, molecular modelling and a preliminary zebrafish model we managed to provide genetic proof for these mutations. By using enrichment analysis from co-expression networks and enzymatic analysis, we saw that the majority of *NARS1* mutations cause a loss of the enzymatic protein by reduced expression and disruption of dimer formation, resulting in abnormal protein synthesis with a possible compensatory effect by other ARSs.

Interestingly, we found out that heterozygous *de-novo* variant such as the c.1600C>T, p.Arg534* mutation, can have a gain-of-function effect that interferes

with normal protein function. A protein lacking the 15 amino acids containing the ATP-binding domain is can escape mRNA decay but also produce a dominant negative effect on the wild-type allele in a zebrafish model, causing cyclopia defects (linked before with microcephaly).

Our data revealed an implication of *NARS1* in paediatric neurodevelopment, in close proximity to other ARSs with similarly severely affected patients, broadening the spectrum of human diseases due to mutations in tRNA synthetase genes. It is of upmost importance that such genes are with time, added to genetic testing panels for children and young adults presenting with NDD, epilepsy, and/or a demyelinating neuropathy.

In chapter 5, I report on a synaptic gene, *SLITRK3*, recently identified as an important transmembrane protein involved in controlling neurite outgrowth. It has never been associated with disease until we identified 3 biallelic or monoallelic mutations in patients presenting with epileptic encephalopathy. The neurological manifestations we report are microcephaly, intellectual disability, seizures, and global developmental delay. By electrophysiologically characterizing mouse-derived hippocampal neurons expressing ST3 plasmids with the mutants, we found that both ST3-C566R and ST3-E606X exhibit impaired development of peripheral and central nervous system through a combination of loss of WT-ST3 protein and a loss of function effect. In these ST3 mutants, the SLITRK3 protein is either truncated or impaired to regulate the inhibitory synapses showing the importance of single critical residues in ST3 C-terminus for GABAergic synapse development and function. In

addition, our results support a conserved role of this transmembrane protein in neurological function and an essential role of ST3-C566R.

Expanding the disease spectrum and/or patient cohorts in known disease-causing genes

Downregulation of SFPQ function has been implicated as a risk factor for human neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), fronto-temporal dementia (FTD) and Alzheimer's Disease (AD). In chapter 3, I show that the recently identified *SFPQ* variant linked to a complex neurological phenotype (characterized by intellectual disability, peripheral neuropathy, vacuolar myopathy, parkinsonian features and neuroradiological anomalies resembling neurodegeneration with brain iron accumulation) can have a deleterious effect on the ability of this mRNA splicing factor protein to maintain normal hindbrain and midbrain separation and normal axonal growth and branching during embryonic development of zebrafish. Given that loss of SFPQ has been studied already across species such as in human tissue, mouse, iPSC-derived neurons, and zebrafish models, we are very interested to further study this gene on both a cellular and model level, in an attempt to find out if associated changes in SFPQ function, can affect the skeletal muscle physiology or iron homeostasis. It is important to assess how biallelic or monoallelic *SFPQ* gene variants are able to contribute to different neurological and degenerative features in humans.

In chapters 4 (*TARS1*) and chapter 5 (*PIGS*), I show that a significant proportion of rare neurodevelopmental diseases are associated with mutations in previously known genes rather than new disease genes. The *TARS1* gene has been linked to

trichothiodystrophy (TTD), or brittle hair syndrome associated with a wide spectrum of clinical features including intellectual disability, decreased fertility, and short stature, while the *PIGS* gene consistently associated with an epileptic-dyskinetic encephalopathy with features of early infantile epileptic encephalopathy, profound disability, and premature death.

Such studies further expand the phenotypic and genotypic spectrum of variants, i.e increasing the cohort of 7 already-published individuals with biallelic pathogenic variants in *PIGS* by 5 new patients, but importantly by reporting on additional neuroimaging findings seen in our patients. This study helps in revealing the first genotype-phenotype associations, which will help clinicians and geneticists in distinguishing *PIGS*-associated disorders from other paediatric neurological syndromes, using techniques such as flow-cytometry analysis.

In the *TARS1* study, we expand the cohort of 2 already-published individuals by 3 new patients with overlapping features to the initial phenotype. Enzymatic and modelling studies on the role of the threonyl-tRNA synthetase enzyme will shed more light on the function of this gene as well as in explaining how variants can have an effect on amino acid charging of tRNA, important for protein translation. We anticipate than when the yeast complementation assays are completed, we will be able to say which of these variants are loss-of-function more confidently as well as further delineate the molecular pathomechanism of *TARS1*-related disorders. The work besides identifying new genes, is also very important to improve management, develop treatments, and encourage the expansion of diagnostic genetic testing to include these genes in diagnostic panels.

The importance of research and clinical practice in disease discovery

For patients with such severe neurological conditions, a molecular diagnosis may have fundamental practical implications. Many of these genetic diseases start very early in life, and for the patient and their family, a diagnostic odyssey begins when the first symptoms appear. Being able to identify the disease-causing variant responsible for their condition and reporting it to them is essential.

When new candidate genes are identified, such as *NFASC* and *NARS1*, it is crucial that their association with disease is validated in a timely manner and incorporated into clinical practice as soon as possible for the benefit of clinical diagnosis. As researchers, we were able to provide extensive evidence of a variant's pathogenicity through functional studies. A collective method in recent years, in the era of next generation gene discovery, is to test the effect of a genetic variant in model systems, including various organisms and/or patient-derived cells. Gene-specific, biochemical and modelling assays can also be used to provide a certain level of supporting evidence. As the medical and scientific communities learn more about the aetiology and the natural history of these genetic disease, we can become more knowledgeable in cases with a poor prognosis in order to discuss with the family the most appropriate end-of-life decisions, or help them with further reproductive choices.

It is also very important that these genes become part of virtual gene panels such as the recently established Genomics England's 'PanelApp' (allowing virtual gene panels to be set up, deposited and queried) or establish these as soon as their association is validated by a functional study. Updating databases such as OMIM is

also a vital tool for the clinical community and one which has proved useful in the case of the *NFASC*-associated disease.

Future directions and the power of model organisms

Since the late 1990s there has been a growing interest and scientific need to use model systems, including various organisms or patient-derived cultured cells, in understanding disease and its course. The complex clinical phenotype of the *NFASC* disease mentioned in this thesis, suggests that the pathology would be difficult to model in flies and the progressive nature of the *NARS1* disease and onset would not be ideal to model in zebrafish. Therefore, developing mouse models is important for rare diseases as it will give an inside track and add to our understanding of common disorders and lead to potential treatments. In addition, mouse CNS neuropathology and peripheral nerve pathology will be important to compare with human tissue.

Future work on a neurofascin mouse model is explained in the main chapter. Such a model could lead to the development of treatments, as it would provide the possibility to rescue nodal function using systems like CRISPR/Cas9 to knock out NFASC and investigate how other nodal genes/ proteins function and interact without it, or how its dysregulation or absence influence the appearance and progression of the demyelinating neuropathy. Upregulation of the other interacting nodal genes such as *CNTN1* and *CASPR1* could also be used to study how the myelination and axogenesis in the mouse model could be affected in a possible rescue experiment.

Understanding how the oligodendrocyte neurofascin regulates both myelin targeting and sheath growth at the same time in individual cells is the next vital path of

exploration. Klingseisen *et al.* (2019) have recently tried to disentangle the functional consequences of cell body myelination and impaired sheath growth, by using a cokbiantion of zebrafish and mouse models. They carried out so far a comparison of neurofascin mutants with disruption to oligodendrocyte and neuronal Caspr and found out that loss of oligodendrocyte neurofascin leads to mistargeting of myelin to cell bodies, without affecting targeting to axons. Its loss also reduces CNS myelination by impairing myelin sheath growth. By using techniques like time-lapse imaging, they revealed that the distinct myelinating processes of individual oligodendrocytes can engage in target selection and sheath growth at the same time and that neurofascin concomitantly regulates targeting and growth. Disruption to neuronal Caspr also impairs myelin sheath growth, likely reflecting its association in an adhesion complex at the axon-glial interface with neurofascin. Caspr does not, however, affect myelin targeting, further indicating that neurofascin independently regulates distinct aspects of CNS myelination by individual oligodendrocytes *in vivo*[291].

Studying a gene such as *NFASC* further after identifying it, is important not only to understand fundamental mechanisms, but also to gain insight into the increasing number of human diseases in which disruption to oligodendrocyte Neurofascin or axonal Caspr are evident.[1; 2; 292-294] The mechanistic dissection of how oligodendrocyte neurofascin, as a single cell adhesion molecule, can regulate two distinct and essential aspects of myelination in the same cell at the same time also represents an important challenge for a future treatment.

In relation to the discovery of the *NARS1* gene, we have since found more families with *NARS1* mutations with (primary) microcephaly, global developmental delay, intellectual disability and varying degree of neurological features. We are hopeful for a mouse model to be established soon at MRC Harwell, since we are interested to understand the role of *NARS1* in wiring the brain in mouse, studying genetic redundancy and signalling mechanisms. We would employ techniques such as proteomics to study the downstream effects of impaired translation. We will also perform behaviour studies such as those designed and used as part of the International Mouse Phenotyping Consortium, to define in detail the phenotype we obtain. Genetic efficiency in higher organisms depends on mechanisms to create multiple functions from single genes. It has been shown that aminoacyl-tRNA synthetase genes identified in this thesis, are used both as enzymes and alternately as regulators of signalling pathways. The wider academic community would benefit from greater understanding of the function of *NARS1* because of the fundamental role it plays both in translation and likely other pathways.

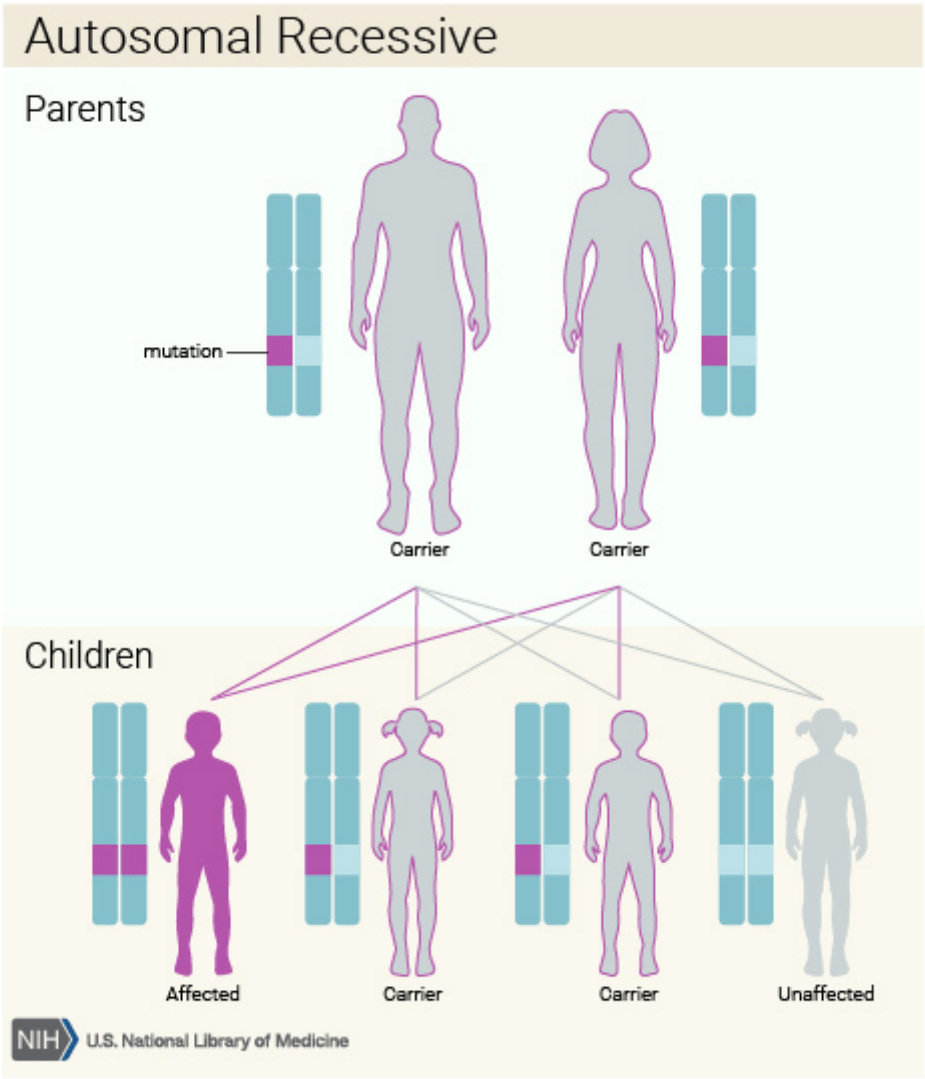
This is a very topical area for the development of potential therapies and other academic groups and industry will be interested once these mouse model have been developed and characterised. Having the patients with these defects on hand for comparison and therapeutic purposes will also be of great significance. This model could lead to the development of treatments to patients with disorders of this overall enzyme defect pathway and in the broader group of neurological conditions. Gene therapy is increasingly being considered as a realistic treatment option for human disorders including a number of successful clinical gene therapy trials. A recent gene therapy trial showed therapeutic benefits in mice with CMT1X, another devastating

inherited neuropathy disorder lacking in medical treatment. This study demonstrated improvement in motor skills and sciatic nerve conduction velocity, and provides impetus to study potential gene therapies in our newly-discovered genes too.[295-297]

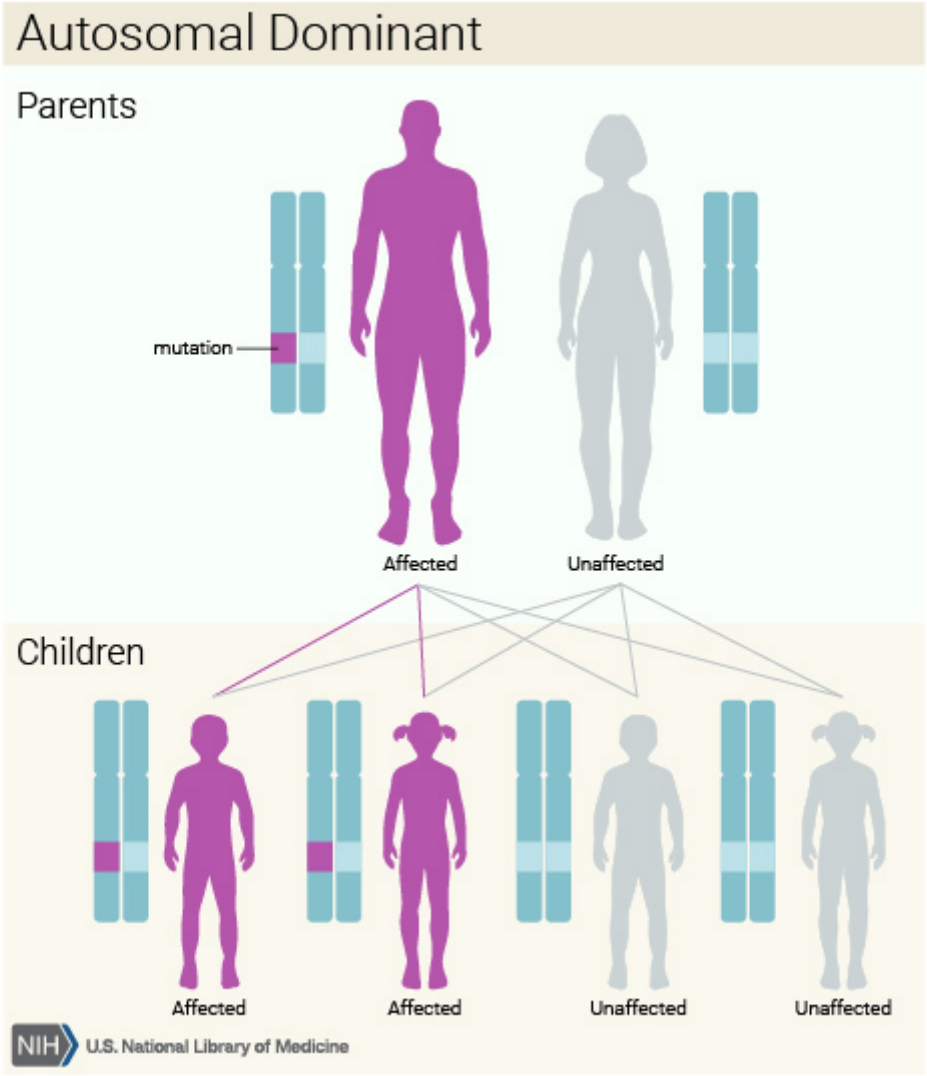
To conclude, my thesis is a documentation of the hard work required for disease-gene discovery that has enabled me and the team to identify disease-mechanisms, as well as define genetic, clinical and biochemical links in an attempt to understand the neuropathology of rare neurological diseases.

Appendix

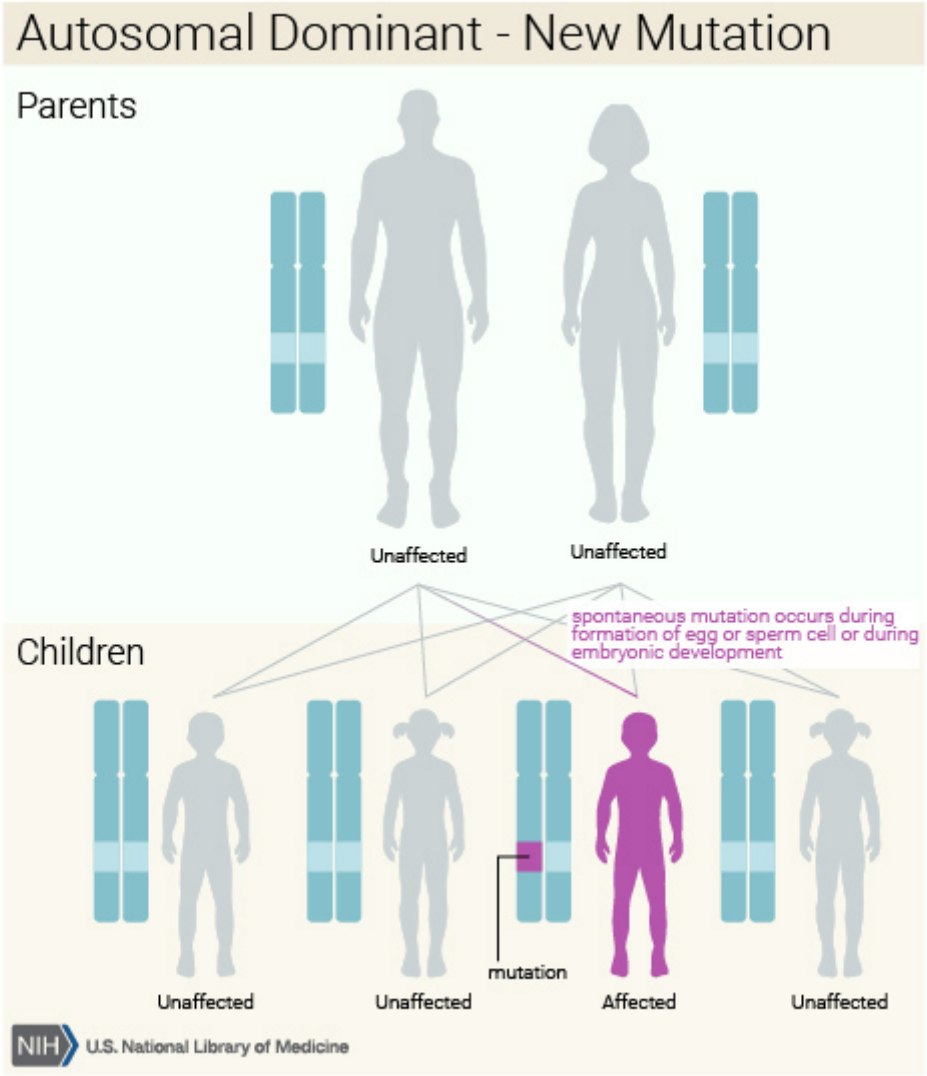
Overview of inheritance modes



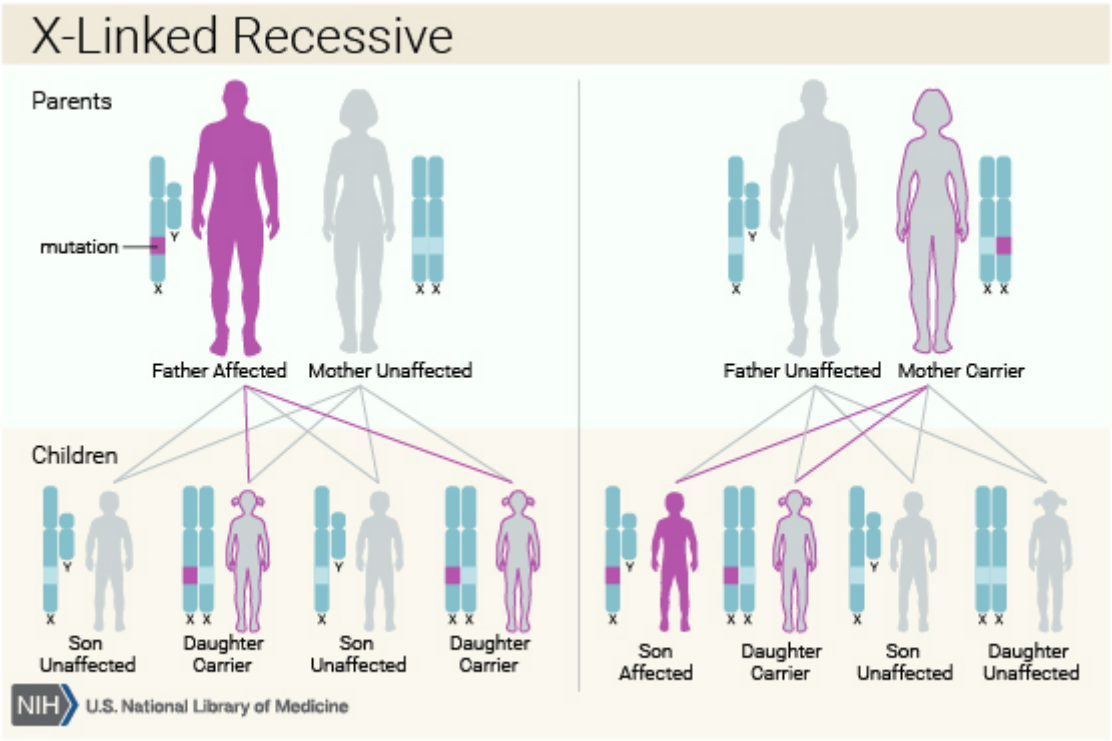
Supplementary Figure 1. Autosomal recessive inheritance. Both copies of the gene in each cell have mutations. The parents of an individual with an autosomal recessive condition each carry one copy of the mutated gene, but they typically do not show signs and symptoms of the condition. Autosomal recessive disorders are typically not seen in every generation of an affected family (from https://ghr.nlm.nih.gov/primer/inheritance/inheritancepatterns).



Supplementary Figure 2. Autosomal dominant inheritance. One mutated copy of the gene in each cell is sufficient for a person to be affected by an autosomal dominant disorder. In some cases, an affected person inherits the condition from an affected parent (from https://ghr.nlm.nih.gov/primer/inheritance/inheritancepatterns).



Supplementary Figure 3. De-novo inheritance. In other cases, the condition may result from a new mutation in the gene and occur in people with no history of the disorder in their family (from https://ghr.nlm.nih.gov/primer/inheritance/inheritancepatterns).



Supplementary Figure 4. X-linked recessive disorders caused by mutations in genes on the X chromosome. In males (who have only one X chromosome), one altered copy of the gene in each cell is sufficient to cause the condition. In females (who have two X chromosomes), a mutation would have to occur in both copies of the gene to cause the disorder. Because it is unlikely that females will have two altered copies of this gene, males are affected by X-linked recessive disorders much more frequently than females. A characteristic of X-linked inheritance is that fathers cannot pass X-linked traits to their sons (no male-to-male transmission) (from https://ghr.nlm.nih.gov/primer/inheritance/inheritancepatterns).

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact *BP4* Missense when only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7* In-frame indels in repeat w/out known function *BP3*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		N ≤ 1/8 if 1 family N ≤ 1/4 if > 1 family	N ≤ 1/16 if 1 family N ≤ 1/8 if > 1 family	N ≤ 1/32 if 1 family N ≤ 1/16 if > 1 family	
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

Supplementary Figure 5. ACMG criteria according to the Association for Clinical Genomic Science (ACGS). This chart organizes each of the criteria by the type of evidence as well as the strength of the criteria for a benign (left side) or pathogenic (right side) assertion. BS, benign strong; BP, benign supporting; FH, family history; LOF, loss of function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong. If a variant does not fulfil criteria using either of these sets (pathogenic or benign), or the evidence for benign and pathogenic is conflicting, the variant defaults to uncertain significance.

GTEx Tissue in which network was constructed	Module Membership	Module Name	Module size	FDR-corrected GO term enrichment p-values	Cell type predictions and associated p-values
Tibial nerve	0.8287	purple	296	ensheathment of neurons GO:0007272 (p-value 0.0132), axon ensheathment GO:0008366 (p-value 0.0132)	Oligodendrocyte-External (p-value 1.588e-07).
Amygdala	0.8741	darkgreen	529	oligodendrocyte differentiation GO:0048709 (p-value 2.61e-06), glial cell differentiation GO:0010001 (p-value 6.64e-06), gliogenesis GO:0042063 (p-value 2.22e-05), regulation of gliogenesis GO:0014013 (p-value 0.000256), glial cell development GO:0021782 (p-value 0.00369)	Oligodendrocytes in Human brain Module (Geschwind,2010) (p-value 4.267e-74). Oligodendrocytes in Cortex (p-value 8.137e-33). Oligodendrocytes, definite (Cahoy, 2008) (p-value 2.02e-14). Oligodendrocytes from conservative data set (Lein, 2007) (p-value 4.81e-06). Oligodendrocyte-External (p-value 1.195e-23).
Anterior Cingulate Cortex	0.7396	turquoise	3916	organelle organization GO:0043933 (p-value 5.35e-10), chemical synaptic transmission GO:0099536 (p-value 6.68e-10), mitochondrion organization GO:0061024 (p-value 7.25e-09), protein ubiquitination GO:0006464 (p-value 1.01e-08), protein modification by small protein conjugation or removal GO:0036211 (p-value 1.49e-08)	Neuron in Human brain Module (Geschwind,2010) (p-value 2.599e-108). Neuron module in Cortex (p-value 6.091e-72). Neuron, pyramidal in network from Sugino/Winden (p-value 1.578e-11). Neuron, definite (Cahoy, 2008) (p-value 3.405e-35). Neuron, probably (Cahoy, 2008) (p-value 2.007e-138). Neurons-Cahoy (p-value 0.006484). Neuron_Pyramidal_CA1-External (p-value 1.343e-12). Neuron_Interneuron-External (p-value 2.68e-15). Neuron_Pyramidal_S1-External (p-value 1.828e-07). Neuron_Dopaminergic_SNigra-External (p-value 3.798e-06). Neuron.In1-External (p-value 0.002993). Neuron.Ex2-External (p-value 0.01224). Neuron.Ex4-External (p-value 0.00388). Neuron.Ex7-External (p-value 0.008385).
Caudate	0.7992	black	945	nervous system development GO:0007399 (p-value 6.02e-13), neurogenesis GO:0022008 (p-value 8.88e-12), ensheathment of	Oligodendrocytes in Human brain Module (Geschwind,2010) (p-value 1.442e-183). Oligodendrocytes in Cortex (p-value 2.687e-97). Oligodendrocytes, definite (Cahoy, 2008) (p-value 1.284e-39). Oligodendrocytes from conservative data set (Lein, 2007) (p-value

				neurons GO:0007272 (p-value 1.87e-11), axon ensheathment GO:0008366 (p-value 1.87e-11), myelination GO:0042552 (p-value 8.78e-11)	3.706e-25). Oligodendrocytes-Cahoy (p-value 1.285e-05). Oligodendrocyte-External (p-value 8.934e-59).
Cerebellar Hemisphere	0.8597	lightcyan1	338	cytoskeleton organization GO:0007010 (p-value 0.0368)	void
Cerebellum	0.6908	black	301	cellular component organization NA (p-value 0.0152), cellular component organization or biogenesis NA (p-value 0.0264)	void
Cortex	0.8084	brown	862	ensheathment of neurons GO:0007272 (p-value 2.8e-09), axon ensheathment GO:0008366 (p-value 2.8e-09), myelination GO:0042552 (p-value 1.45e-08), neurogenesis GO:0022008 (p-value 3.35e-06), oligodendrocyte differentiation GO:0048709 (p-value 6.76e-06)	Oligodendrocytes in Human brain Module (Geschwind,2010) (p-value 3.216e-141). Oligodendrocytes in Cortex (p-value 3.19e-72). Oligodendrocytes, definite (Cahoy, 2008) (p-value 7.64e-39). Oligodendrocytes from conservative data set (Lein, 2007) (p-value 1.04e-16). Oligodendrocytes-Cahoy (p-value 1.207e-05). Oligodendrocyte-External (p-value 3.665e-46).
Frontal Cortex	0.7939	red	1464	covalent chromatin modification GO:0016569 (p-value 0.000148), organelle organization GO:0006996 (p-value 0.000699), nucleic acid-templated transcription GO:0097659 (p-value 0.00114), nucleobase-containing compound biosynthetic process GO:0034654 (p-value 0.00129), aromatic compound biosynthetic	void

				process GO:0019438 (p-value 0.0015)	
Hippocampus	0.8811	antiquewhite4	339	void	Oligodendrocytes in Human brain Module (Geschwind,2010) (p-value 6.053e-07). Oligodendrocytes in Cortex (p-value 0.008201). Oligodendrocyte-External (p-value 5.488e-10).
Hypothalamus	0.8889	darkred	1218	RNA processing GO:0006396 (p-value 1.25e-10), mRNA processing GO:0006397 (p-value 2.55e-10), mRNA metabolic process GO:0016071 (p-value 7.34e-08), RNA splicing, via transesterification reactions GO:0000375 (p-value 2.69e-07), RNA splicing, via transesterification reactions with bulged adenosine as nucleophile GO:0000377 (p-value 5.95e-07)	void
Nucleus Accumbens	0.8383	sienna3	789	regulation of gene expression GO:0010468 (p-value 9.02e-11), regulation of RNA metabolic process GO:0051252 (p-value 6.55e-10), RNA biosynthetic process GO:0032774 (p-value 7.16e-10), nucleic acid-templated transcription GO:0097659 (p-value 2.34e-09), regulation of nucleobase-containing compound metabolic process GO:0019219 (p-value 2.8e-09)	void
Putamen	0.8211	pink	824	cell cycle process GO:0022402 (p-value 0.00325), organelle fission	void

				GO:0048285 (p-value 0.00882), mitotic cell cycle process GO:1903047 (p-value 0.0139), mitotic nuclear division GO:0007067 (p-value 0.0142), nuclear division GO:0000280 (p-value 0.0152)	
Spinal cord	0.844	plum	556	ensheathment of neurons GO:0007272 (p-value 0.00018), axon ensheathment GO:0008366 (p-value 0.00018), myelination GO:0042552 (p-value 0.00102)	Oligodendrocytes in Human brain Module (Geschwind,2010) (p-value 2.535e-33). Oligodendrocytes in Cortex (p-value 7.784e-13). Oligodendrocytes from conservative data set (Lein, 2007) (p-value 0.01577). Oligodendrocyte-External (p-value 0.001345).
Substantia nigra	0.8884	bisque4	508	carbohydrate phosphorylation GO:0046835 (p-value 0.0145)	void

Supplementary Table 1. Gene co-expression analysis using all 13 human brain region networks sampled by the Genotype Tissue Expression Consortium.

Supplementary Clinical Data for Chapter 4.4 (clinical data was provided by Dr. Emer O'Connor)

Family 1: Individual 1: *NARS1 de-novo* mutation, c.1600C>T, p.Arg534*

The proband was a Dutch female born to healthy non-consanguineous parents. She had an uneventful perinatal course and was born with a normal weight for gestational age. Feeding difficulties were immediately apparent, requiring tube feeding in neonatal period. As an infant, she had global development delay (GDD), failing to meet developmental milestones in multiple areas of functioning. In terms of motor milestones, sitting was achieved aged 18 months and she began walking aged 36 months. She had severe delay in language development, speaking first words aged 2 year, and at follow-up aged 17 years, her vocabulary is limited to approximately 10 words. She was also found to have progressive microcephaly with an occipital frontal circumference (OFC) of 45.8 cm (2^{nd} percentile, -2.1 SD) at 3.5 years and 49 cm ($<1^{st}$ percentile, -5 SD) at 16 years. She had an MRI aged 4 years, which was otherwise unremarkable. Recurrent febrile seizures were reported aged of 11, 16, 23 and 33 months. As she grew older, she continued to exhibit signs of severe intellectual disability with limitations across all adaptive domains. Her examination was notable for a number of dysmorphic features including brachycephaly, deep-set eyes, upslanting palpebral fissures, short philtrum, long slender fingers and persistent fetal finger pads. She had bilateral pes cavus requiring surgical correction and a unilateral foot-drop. She was unable to fully co-operate in a neurological examination, but there was evidence of wasting distally in the upper and lower limbs, she also appeared ataxic on mobilising and had absent ankle reflexes bilaterally.

Family 2: Individual 2: *NARS1 de-novo* mutation, c.1600C>T, p.Arg534*

A Dutch female born to non-consanguineous parents. She presented with global developmental delay with failure to meet age-appropriate milestones, first sitting aged 16 months old and walking at 26 months. She spoke her first words aged 2 years and 6 months and continued to show profound speech delay. Now, as an adult, she speaks only a few words with a nasal pitch and has severe intellectual disability. Microcephaly was confirmed on examination with an OFC measuring 46.8 cm (<1st percentile, -4.3 SD) aged 8 years and 11 months. She has epilepsy, experiencing her first GTC seizure aged 3 years, with epileptiform discharges confirmed on EEG in the posterior temporal regions. On examination, she has several dysmorphic features including mild upslanting palpebral fissures, wide set teeth, a broad jaw, clinodactyly and marked thoracic kyphosis. She has severe bilateral foot drop with atrophy of the lower legs and intrinsic foot muscles. She is hypotonic with muscle weakness, which is more pronounced distally. Vibration and sensation are impaired to the level of the ankles. Her coordination is grossly normal but she has a slight intention tremor in the right upper limb. She has had several investigations including an MRI brain, aged 9 years, which was reported as normal. Nerve conduction studies confirmed a demyelinating polyneuropathy and muscle ultrasound indicated hyperechogenecity distally in lower limbs bilaterally.

Family 3: Individual 3: *NARS1 de-novo* mutation, c.1600C>T, p.Arg534*

The proband was a Dutch male born to healthy non-consanguineous parents. His pre-natal course was remarkable for inter-uterine growth retardation and oligohydramnios. At birth, he was found to have a low weight and OFC for gestational age. GDD was evident from early infancy as he failed to reach several

milestones including gross motor parameters, walking for the first time aged 30 months. In particular, his speech was severely delayed. Now, aged 10 years his language is limited to short sentences with a notably nasal quality. Dysmorphic features were evident on follow-up examination including medial eyebrow flare, a short upturned nose, retrognathia and clinodactyly of his fifth fingers. He was notably dysarthric and had a bilateral tremor with occasional myoclonus. He was hypertonic with clonus at the knee and ankle. Power was reduced with weakness in a pyramidal distribution in the upper and lower limbs. He was hyperreflexic with upgoing plantar reflexes bilaterally.

Family 4: Individual 4: *NARS1* de- novo mutation c.1600C>T, p.Arg534*

Europian male born to healthy non-consanguineous parents. He presented with his first seizure aged 4 months old, suffering from both partial and myoclonic seizures with associated EEG abnormalities. In addition, he has a dilated aortic root, which was diagnosed in infancy. He suffered from GDD. Gross motor skills were severely delayed, first sitting aged 16 months and never acquiring the ability to walk unaided. His speech was also profoundly delayed and is now limited to 1-2 words at the age of 13 years. On examination, he had a dysmorphic appearance with arachnodactyly, pectus excavatum, dolichostenomelia, long palpebral fissures and hypertelorism. He exhibited stereotypies with hand and mouth repetitive movements. He had microcephaly with an OFC of 49cm aged 13 years old (<1st percentile, -4.7 SD). His neurological examination was notable for severe spasticity in upper and lower limbs and impaired co-ordination.

Family 5: Individual 5: *NARS1 de-novo* mutation c.1600C>T, p.Arg534*

Female born to healthy non-consanguineous parents. Characterised by severe global developmental delay, with profound speech delay. Now aged 16 using she has only 1-2 word phrases. She first walked age 3 years. She had feeding difficulties in infancy requiring G-tube feeding. In terms of epilepsy, she had an initial seizure at 3 months of age continued to have both partial and generalized seizures throughout her childhood, which was well managed with levetiracetam. An EEG at 15 years was consistent with chronic static encephalopathy. Her examination was notable for microcephaly with a sloping forehead, slanting eyes, low set ears with fleshy helices, widely spaced teeth and small hands with tapered fingers. She walked with a broad-based ataxic gait. There was wasting distally especially involving the extensor digitorum brevis. She was hypotonic in upper and lower limbs with distal weakness and hyporeflexia. She had impaired sensation to pinprick in the upper and lower limbs. Nerve conduction studies showed findings consistent with a demyelinating neuropathy.

Family 6: Individual 6: *NARS1 de-novo* mutation c.1600C>T, p.Arg534*

Male proband born to healthy non-consanguineous parents. He presented in infancy with neurodevelopmental delay (NDD), walking for the first time aged 23 months. His speech development was severely delayed and, now aged 8 years, he has a vocabulary limited to approximately 20 words. He began having seizures at 11 months, which were initially considered febrile seizures; however, these occurred recurrently throughout childhood and were classified as atypical febrile seizure. At this time, he underwent an EEG which failed to capture epileptiform activity. On examination he had microcephaly (<1st percentile, -3.25SD) and several dysmorphic

features including a long philtrum, a thin upper lip, an everted lower lip, a wide mouth, midface hypoplasia and low set ears with overfolded helices. He also had syndactyly of his toes and short first toe bilaterally. His physical examination was notable for a broad-based ataxic gait and increased tone in the lower limbs bilaterally with hyperreflexia and clonus. MRI brain showed mild cortical atrophy with enlargement of the CSF space at 8 months however subsequent imaging aged 3 years was unremarkable for structural abnormalities.

Family 7: Individual 7: *NARS1 de-novo* mutation c.1525C>T, p. Gly509Ser

The proband was a British female born to non-consanguineous parents. Of note, exome sequencing also revealed a complex X chromosome rearrangement. She first presented at 9 months with focal seizures with secondary generalisation. She subsequently underwent EEG, which failed to capture epileptiform discharges. She had GDD, failing to reach a number of developmental milestones, resitting for the first time aged 12 months and walking aged 27 months with severe delays in speech. She persisted to have learning difficulties throughout her childhood and suffered with chronic constipation. On examination, she was normocephalic. Dysmorphic features included broad forehead, large ears, tented upper lip, long and slender fingers, slender feet, hypermetropia and a unilateral convergent squint. She was hypotonic with brisk DTR and down-going plantars. She unexpectedly died while sleeping aged 10 years 8 months. Autopsy showed a normally formed brain with focal calcification of basal ganglia and dentate nucleus.

Family 8: Individual 8: *NARS1 de- novo* mutation c.965 G>T, p.Arg322Leu

The proband is a 15-year-old male of mixed race (German-Irish / English-Native American/ Russian-Polish heritage) born to healthy parents. In infancy, he had feeding difficulties and was slow to reach motor milestones, sitting aged 8 months and walking at 35 months. His speech was profoundly delayed speaking first words aged 8-9 years old. He has intractable epilepsy characterised by both absence and myoclonic seizures, confirmed on EEG with slow wave bursts with spike & poly-spike bursts. On examination, dysmorphic features included a prominent nose and broad forehead. He is dysarthric with limited speech, which mostly consists of repetitive phrases. He had increased tone in the lower limbs bilaterally and hyperreflexia throughout. Vibration and sensation were intact. Co-ordination was impaired with both appendicular and axial ataxia. MRI was unremarkable for structural abnormalities.

Family 9: Individual 9 and 10: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

The proband, individual 9, is a 33-year-old male of Indian descent, born to consanguineous parents. He had a normal prenatal course and uncomplicated birth. He failed to meet developmental milestones in infancy, with severe delays in speech and fine motor skills. He began having generalised tonic-clonic seizures aged one year, which were poorly controlled with anti-epileptic agents and persisted into adulthood. Contractures in his lower limbs required him to undergo a tendon lengthening procedure aged 14 years. Additionally, he has severe sensory and motor neuropathy with NCV showing complete absence of sensory action potentials in the upper limbs, and absent motor action potentials in the upper and lower limbs.

On examination, he is microcephalic and has bilateral foot drop with a left sided foot deformity consistent with a Charcot joint. He wears AFOs to enable ambulation and has a broad-based ataxic gait. There are contractures of his fingers bilaterally and evidence of muscle atrophy distally. He is hypotonic with reduced power proximally (3/5) with a pronounced weakness distally (2/5). He is areflexic and has impaired sensation to pinprick to the level of elbows and ankles in the upper and lower limbs respectively. His younger brother, individual 10, is similarly affected. He also had an unremarkable prenatal course and birth but failed to meet expected developmental milestones. As an older child, he had learning difficulties requiring special schooling. He has no history of seizures. He has scoliosis, a broad-based gait and difficulty walking. Nerve conduction velocities revealed a demyelinating neuropathy in upper and lower limbs.

Family 10: Individual 11: *NARS1* Homozygous mutation c.1633C>T, p. Arg545Cys

The proband, is an 8-year-old male born to consanguineous parents from Pakistan. As an infant, he initially had feeding difficulties, choking and regurgitating with feeds. He failed to meet developmental milestones with severe delays in gross motor skills, sitting at 12 months and walking at 24 months. Speech was also severely delayed and now, aged eight, his vocabulary is limited to a few sentences. On examination, he had severe microcephaly with an OFC of 46.5cm (<1p, -4.2 SD). He had fifth finger clinodactyly bilaterally and toe syndactyly of the right foot. Neurological examination was limited in a setting of intellectual disability. There was no evidence of wasting. He had a broad-based gait and poor balance with difficulty running.

Patellar reflexes were 3+ bilaterally; otherwise reflexes were normal with down-going plantars. His MRI was unremarkable for structural abnormalities.

Family 11: Individuals 12-15: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

The proband is 17-year-old female born to Pakistani parents. She has severe intellectual disability across all adaptive domains. As a child, she was slow to meet several developmental milestones, sitting for the first time at 12 months and walking at 20 months. She also had difficulty feeding, choking regularly. She has microcephaly with and OFC of 49.5 cm (<1p,-4.5 SD). She has no history of seizures. Her physical examination is notable for reduced power (3/5) in all muscle groups with and an ataxic gait. She has no evidence of wasting or impaired sensation. She has two siblings, an older brother and sister and a female cousin who are similarly affected with GDD in childhood and severe intellectual disability. Additionally, her older brother and sister also have epilepsy characterised by generalised tonic clonic seizures and her eldest sibling is the most severely affected with profound intellectual disability.

Family 12: Individuals 16-17: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

The proband is a male born to consanguineous parents was reviewed aged 8 years. He has severe intellectual disability with a history of GDD, including severe delay in meeting motor milestones, walking for the first time age 6.5 years. He also had severe speech delay and at the time of follow-up had a limited vocabulary and difficulty forming sentences. His neurological examination was notable for hypotonia

and lower limb weakness. His brother, now aged 21 years, is also affected with severe intellectual disability, wasting, and weakness in lower limbs.

Family 13: Individuals 18-19: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

Two siblings born to consanguineous parents from Pakistan. The proband is 6 months old male who presented in the neonatal period with microcephaly, seizures and failure to thrive. His older brother also has epilepsy, microcephaly and global developmental delay. He walked age 3 years. His speech is severely delayed and, at follow-up aged 6.5 years, he only speaks a few words and is unable to form sentences.

Family 14: Individuals 20: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

The proband is a 16-year-old boy born to consanguineous parents from Pakistan. He was born following a normal pregnancy and perinatal course, but began to exhibit signs of moderate global developmental delay in infancy. He sat at 7 months, began walking at 2 years of age and had moderate delay in speech development. He had learning disabilities in school and persisted to have moderate intellectual disability. His examination was notable for bilateral pes cavus, weakness and wasting distally in upper and lower limbs, absent deep tendon reflexes and impaired sensation. Nerve conduction velocities revealed severe sensorimotor polyneuropathy with primary axonal degeneration. His MRI Brain was unremarkable.

Family 15: Individuals 21-23: *NARS1* Homozygous mutation c.1633C>T, p.Arg545Cys

This family consists of eight siblings born to consanguineous Pakistani parents. Four siblings were affected with GDD, microcephaly and seizures. The eldest, a female, was reviewed aged 30 years. She presented with severe GDD in childhood, sitting for the first time at 4 years of age and walking first steps at 10 years. She did not speak at all until she was 4.5 years old and now has limited speech with difficulties communicating. At 4.5 years of age she fell downstairs, after which, she began having left sided focal seizures, which were controlled with sodium valproate. Her seizures stopped after 2 years, however she had persistent left sided weakness which was attributed to the fall. On examination, she had bilateral foot drop and several dysmorphic characteristics including; large dysplastic ears, a large nose with broad nares, a large mouth with widely spaced teeth, and syndactyly of the second and third toes bilaterally. She had normal tone with reduced power in lower limbs bilaterally particularly on dorsiflexion (1/5). She had an ataxic gait with absent ankle reflexes. She was unable to participate in sensory examination.

Her younger brother, who is now 16, is similarly affected but has less severe intellectual disability. He also has microcephaly and had delayed developmental milestones as child, sitting at 1 years of age, walking aged 2 years and speaking first words at 3 years of age. His speech remains limited, he is unable to speak in full sentences, and he has a tendency towards aggressive behaviour. He also had focal seizures in infancy, however they have now resolved and his last witnessed seizure occurred aged 2 years. Two other siblings, also affected, died in childhood. A girl, who was reviewed aged 13 years, also presented with GDD and microcephaly in infancy. She had severe intellectual disability and was unable to feed herself

independently. She also had a history of delayed milestones in childhood, sitting aged two and walking aged three, and had poor language abilities and spoke only a few words. She began having generalised tonic-clonic seizures aged 4 years which were treated with sodium valproate but poorly controlled. She died at the age of 16 years due to sepsis secondary to burns which she obtained following a fall into a stove. Her brother also died aged 6 years of age. He was unable to sit, stand or speak and had severe and seizures which commenced when he was 2 months old. His death was attributed to respiratory distress secondary to unrelated respiratory complications. Besides these siblings, these individuals also had six first degree relatives (not shown) that had GDD and died in early childhood.

Family 16: Individuals 24-25. *NARS1* siblings with homozygous mutation c.32G>C, p.Arg11Pro

Two affected siblings born to parents from Kosovo, who were not known to be related. The couple's first child (a girl) was unaffected. Their second child, a boy, was born at full term following an uneventful pregnancy. He had a normal postnatal course and birth parameters. No abnormalities were detected until he was 5 months old when he presented with generalised tonic-clonic seizures. Moreover, his motor development deteriorated with progressive muscle spasticity. CT imaging of his brain revealed widening of bifrontal subarachnoid spaces as well as small subdural hygroma. Laboratory work-up including CSF analysis was unhelpful. Brain MRI at the age of 10 months revealed prominent cerebral atrophy and delayed myelination. During the following year, he was treated with phenobarbital, levetiracetam and baclofen to control seizures and spasticity. Progressive microcephaly was noted OFC 43 cm (<1st percentile, -4 SD). Motor and cognitive milestones were not

reached (e.g. no crawling, no sitting without support, no development of language, etc). Follow-up brain MRI at the age of 1 ½ years showed severe atrophy with no progression of myelination. At the age of 2 ½ year, the boy deteriorated with poor feeding, vomiting and dehydration. An MRI brain demonstrated basilar thrombosis and associated infarction in the cerebellum, pons and midbrain. His neurological function deteriorated further and he died aged 4 years in a palliative care setting with severe aspiration pneumonia. The third born, a girl, had normal pre- and postnatal development. She presented aged 8 months with status epilepticus. Physical examination revealed microcephaly, developmental delay and moderate spasticity of the extremities. Brain MRI showed comparable findings to her brother with severely delayed myelination. Follow-up brain MRI at the age of 2 years revealed progressive brain atrophy and arrest of myelination.

Family 17: Individual 26. *NARS1* siblings with homozygous c.50C>T, p.Thr17Met

The proband is a female born to Libyan parents who were first cousins. She was reviewed aged 7 years, at which point was found to have severe microcephaly 42cm (<1st percentile, -7.7 SD). She has severe intellectual disability. She presented with generalised tonic-clonic seizure at 6 months and was diagnosed with GDD in infancy failing to meet age-appropriate milestones. She has 2 additional siblings who were similarly affected but passed away in childhood.

Family 18: Individuals 27 and 28. *NARS1* siblings with compound heterozygous c.1049T>C, c.1264G>A, p.Leu350Pro, p.Ala422Thr

The proband is a 15-year-old female born to healthy non-consanguineous German parents. She presented in childhood with neurodevelopmental delay and progressive microcephaly. She failed to reach appropriate motor milestones sitting for the first time age 3 years. Speech developmental was also severely delayed and now aged 15 years her vocabulary is limited to a few words and she is unable to communicate effectively. She persisted to have severe intellectual disability and experiences frequent episodes of inappropriate laughter. She had her first generalised tonic-clonic seizures at approximately 4 years of age, at which point she was documented as having an abnormal EEG. Her neurological examination was notable for an ataxic gait. She had muscle atrophy which is more pronounced in her lower limbs. She is hypotonic with reduced power distally and hyporeflexic in the lower limbs with upgoing plantars bilaterally. Sensation to sharp touch was impaired. A MRI brain ruled out any structural abnormalities and nerve conduction velocities confirmed demyelinating peripheral neuropathy. Her sister who is now 21 years of age and is clinically similar to her sister with severe intellectual disability, microcephaly and demyelinating peripheral neuropathy. However, she has never had seizures.

Family 19: Individual 29-30. *NARS1* siblings with compound heterozygous variants c.1067A>C, c.203dupA, p.Asp356Ala, p.Met69Aspfs*

The proband is a male born to non-consanguineous Turkish parents. He was born with a low weight (-2.38 SD) and height (-3.76 SD) for gestational age. He failed to meet age-appropriate developmental milestones in terms of motor skills and did not walk until 3 years of age. His speech was also severely delayed speaking his first words aged 4 years. He began having generalised tonic clonic seizures aged 6 years at which point he underwent an MRI brain, which revealed thickening of gyri.

Examination at follow up at 14 years of age revealed severe microcephaly 49.2 cm (<1st percentile, -3.4 SD). His sister, also affected, is clinically identical with GDD and epilepsy.

Family 20: Individual 31. *NARS1* compound heterozygous variants c.268C>T, c.394G>T, p.Arg90*, p.Gly132Cys

The proband is an 8-year-old Canadian female born to healthy non-consanguineous parents. She was born at 37 weeks and had feeding difficulties from the beginning with associated failure to thrive. Developmental milestones were profoundly delayed. She did not walk until 6 years and 10 months and did not speak until she was 5 years old. At follow-up aged 8 years, her speech was limited to 3 words and she was unable to form sentences. From 1 to 3 years of age, she experienced five generalised tonic-clonic seizures. She is now seizure free and does not require anti-epileptic drugs. On examination, she was small for age with a weight of 11kg (-5.6 SD) and height of 95.3 cm (-7.6 SD). She has severe microcephaly (-3SD) and several dysmorphic features including hypotelorism, deep set eyes, a prominent nasal bridge and thin upper lip with smooth philtrum. Skeletal abnormalities include a right sided hip dysplasia and bilateral varus deformities requiring de-rotational osteotomies. Due to severe intellectual disability, she struggled to co-operate with the neurological examination, however she appeared to be hypotonic and have impaired co-ordination. An MRI brain showed microcephaly with a thin corpus callosum and decreased white matter volume throughout.

Family 21: Individual 32. NARS compound heterozygous c.1376 C>T, c.178 A>G, p.Thr459Ile, p.Lys60Glu

The proband is a 3-year-old male born in the USA to healthy non-consanguineous parents. He is microcephalic on examination with a head circumference of 43.5cm (-5 SD). His history is notable for global developmental delay characterised by severe language delays and delays in gross motor milestones, sitting at 10 months and starting to walk at 3 years. He now walks with an ataxic gait and has evidence of spasticity on examination. He also has epilepsy and began having generalised tonic clonic seizures in infancy. MRI brain revealed a small arachnoid cyst involving the right middle cranial fossa but was otherwise within normal limits.

Gene	Locus	Location of Protein	Mode	Disease Phenotype(s)	Clinical severity	Ref
AARS1	16q22	Cytoplasm	AD	Charcot Marie Tooth disease type 2N	Mild	[298-302]
			AR	Epileptic encephalopathy, early infantile, 29	Moderate	
AARS2	6p21.1	Mitochondria	AR	Mitochondrial Infantile Cardiomyopathy	Moderate	[303-307]
				Leukoencephalopathy with ovarian failure	Moderate	
CARS1	11p15.4	Cytoplasm	AR	Microcephaly, DD, brittle hair and nails	Moderate	[308]
CARS2	13q34	Mitochondria	AR	Epilepsy encephalopathy, myoclonic epilepsy	Moderate	[309-312]
				Combined oxidative phosphorylation deficiency 27		
DARS1	2q21.3	Cytoplasm	AR	Hypomyelination, brainstem, spinal cord and leg spasticity	Moderate	[313-315]
DARS2	1q25.1	Mitochondria	AR	Hypomyelination, brainstem, spinal cord, elevated lactate	Severe	[316-325]
EPRS1	1q41	Cytoplasm	AR	Leukodystrophy, hypomyelinating, 15	Moderate	[326]
EARS2	16p12.2	Mitochondria	AR	Leukoencephalopathy and high lactate	Severe	[327-332]
				Combined oxidative phosphorylation deficiency 12		
FARSA	19p13.2	Cytoplasm	AR	Rajab interstitial lung disease with brain calcifications	Moderate	[220]
FARSB	2q36.1	Cytoplasm	AR	Rajab interstitial lung disease with brain calcifications	Moderate	[333; 334]
FARS2	6p25.1	Mitochondria	AR	Combined oxidative phosphorylation deficiency 14	Moderate	[335-340]
				Spastic paraplegia 77		
GARS1	7p15	Cytoplasm & Mitochondria	AD	Charcot Marie Tooth disease type 2D	Mild	[221; 299; 341-356]
			AR	Distal SMA type V, myalgia, cardiomyopathy		
HARS1	5q31.3	Cytoplasm	AD	Charcot Marie Tooth disease type 2W	Mild	[357; 358]
			AR	Usher Syndrome 3B	Moderate	
HARS2	5q31.3	Mitochondria	AR	Perrault Syndrome 2	Mild	[359; 360]

IARS1	9q22.31	Cytoplasm	AR	ID, GR, muscular hypotonia, hepatopathy, cholestasis	Severe	[361]
IARS2	1q41	Mitochondria	AR	Cataracts, GH deficiency, deaf, neuropathy, bone dysplasia	Severe	[362-364]
				Leigh syndrome		
KARS1	16q23.1	Cytoplasm & Mitochondria	AR	Intermediate Charcot Marie Tooth disease type B	Moderate	[365-368]
				Autosomal recessive deafness-89		
				Visual impairment, microcephaly, DD, seizures		
				Leukoencephalopathy		
LARS1	5q32	Cytoplasm	AR	Infantile hepatopathy	Severe	[369]
LARS2	3p21.31	Mitochondria	AR	Perrault syndrome 4	Mild	[370]
				Hydrops, lactic acidosis and sideroblastic anaemia		
MARS1	12q13.3	Cytoplasm	AD	Charcot Marie Tooth disease type 2U	Mild	[371-374]
			AR	Interstitial lung and liver disease	Moderate	
MARS2	2q33.1	Mitochondria	AR	DD, sensorineural hearing loss, Spastic ataxia 3	Mild	[375; 376]
				Combined oxidative phosphorylation deficiency 25	Moderate	
NARS2	11q14.1	Mitochondria	AR	Alpers, Leigh syndrome, DD, ID, epilepsy, myopathy	Severe	[236-238; 377-379]
				Combined oxidative phosphorylation deficiency 24		
PARS2	3p21.31	Mitochondria	AR	Infantile-onset developmental delay and epilepsy	Moderate	[237]
				Alpers syndrome		
QARS1	3p21.31	Cytoplasm	AR	Microcephaly, seizures, DD, cerebral cerebellar atrophy	Severe	[380]
RARS1	5q34	Cytoplasm	AR	hypomyelinating leukodystrophy 9	Severe	[381-383]

RARS2	6q16.1	Mitochondria	AR	Epileptic encephalopathy, lactic acidosis neurological symptoms, pontocerebellar hypoplasia 6	Severe	384-386
SARS1	1p13.3	Cytoplasm	AR	Ataxia, weakness, ID, microcephaly, speech impaired	Moderate	387
SARS2	19q13.2	Mitochondria	AR	Hyperuricemia, pulmonary HT, renal failure, alkalosis	Moderate	388
TARS1	5p13.3	Cytoplasm	AR	Trichothiodystrophy, ichthyosis, ID, decreased fertility	Moderate	242
TARS2	1q21.2	Mitochondria	AR	Axial hypotonia and severe psychomotor delay	Mild	389; 390
				Combined oxidative phosphorylation deficiency 21		
VARS1	6p21.33	Cytoplasm	AR	Severe DD, microcephaly, seizures, cortical atrophy	Moderate	216-219
VARS2	6p21.33	Mitochondria	AR	Microcephaly, epilepsy, encephalocardiomyopathy	Moderate	389; 391-394
				Combined oxidative phosphorylation deficiency 20		
WARS1	14q32.2	Cytoplasm	AD	Distal hereditary motor neuropathy	Mild	395
WARS2	1p12	Mitochondria	AR	Ataxia, weakness, microcephaly, speech impaired, ID	Moderate	387; 396-398
YARS1	1p35.1	Cytoplasm	AD	Dominant-intermediate Charcot Marie Tooth disease	Mild	223; 392
			AR	Multisystem disease, DD, failure to thrive	Moderate	
YARS2	12p11.2	Mitochondria	AR	myopathy, lactic acidosis, and sideroblastic anaemia	Moderate	399-401

Supplementary Table 2. ARS gene implicated in human disease. *ARS* gene implicated in 36 human diseases from a total of 37 genes. The first letter of the ARSs' gene name corresponds to the amino acid specificity of the corresponding ARS (based on the amino acid one-letter code) and 2 indicates that the gene encodes a mitochondrial-restricted isoform [208-210]. Clinical phenotype is bases upon

publications. Key: ID = intellectual disability, GH = growth hormone, GR = growth retardation, DD = developmental delay, AR = autosomal recessive, AD = autosomal dominant, HT = hypertension, SMA = spinal muscular atropy

Supplementary Table 3. Detailed clinical features of individuals with *NARS1 de-novo* heterozygous mutations. Families 1-8.								
Family	1	2	3	4	5	6	7	8
Individual	1	2	3	4	5	6	7	8
Inheritance	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous	*de-novo* Heterozygous
Variant	c.1600C>T p.Arg534*	c.1600C>T p.Arg534*	c.1600C>T p.Arg534*	c.1600C>T p.Arg534*	c.1600C>T p.Arg534*	c.1600C>T p.Arg534*	c.1525C>T p.Gly509Ser	c.965 G>T p.Arg322Leu
Origin	Dutch	Dutch	Dutch	European	Hispanic, European	European	UK	European
Consanguinity	N	N	N	N	N	N	N	N
Gender / Age at follow-up	F/ 17	F / 22	M / 10	M /13	F / 16	M / 5y 8m	F / 2y 10m	M /15
Occipital Frontal Circumference (OFC)								
Birth	45.8cm <2p *(-2.1SD)*	NA	NA	33cm (10p)	NA	NA	NA	NA
Follow –up Percentile / SD	49cm <1p *(-5SD)*	46.8cm <1p*(-4.3 SD)*	50cm <1p *(-2.2 SD)*	49cm <1p *(-4.7SD)*	49.3cm at 13 y <1p *(-4.5 SD)*	47cm <1p *(-3.2SD)*	50 cm 86p	54cm 27p
Microcephaly	Y	Y	Y	Y	Y	Y	N	N
Developmental Delay								
GDD	Y	Y	Y	Y	Y	Y	Y	Y
Sitting (Months)	18	16	NA	16	11	NA	12	8
Walking (Years)	3	2y 2 m	2y 6m	N	3	23	2y 3m	2y 11m
Language	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed
Neurological Findings								

Intellectual Disability	Severe	Severe	Severe	Profound	Profound	Severe	Severe	Severe
Seizures	GTC	GTC	NA	Myoclonic / Partial	GTC / Partial	GTC	GTC	Partial/Myoclonic
Peripheral Neuropathy	Y	Y	NA	N	Y	N	NA	NA
Ataxia	Y	Y	Y	Y	Y	Y	NA	Y
Imaging	NAD	NAD	NAD	NA	NA	Mild atrophy, CSF space enlargement	NA	NAD
Clinical Features								
Dysmorphic Features	Upslanting palpebral fissures Pes-cavus	Clinodactyly Upslanting palpebral fissures Thoracic Kyphosis Wide spaced teeth	Clinodactyly Retrognathia	Upslanting palpebral fissures Hypertelorism Arachnodactyly, Pectus Excavatum	Upslanting palpebral fissures Wide spaced teeth Low set ears Fleshy Helices	Low set ears Overfolded Helices Syndactyly	Large ears Long slender fingers	Broad Forehead
Tone	NAD	Reduced	Increased	Increased	Reduced	Increased	Reduced	Increased
Power	Reduced	Reduced	Reduced	NAD	Reduced	Reduced	NA	NAD
Sensation	NA	Reduced	NA	NAD	Reduced	NAD	NA	NA
Co-ordination	Ataxic gait	Unilateral Intention Tremor	Dysarthria	Ataxic gait	Ataxic gait	Ataxic gait	NA	Ataxia
Reflexes	Reduced	Reduced	Increased	Increased	Reduced	NAD	Increased	Increased
Other	N	N	Tremor /Myoclonus	Stereotypies	N	N	N	Stereotypies

Y = Yes, N= No, M= Male, F= Female, NA= Not Available, p = percentile, SD = standard deviation, NAD = No abnormalities detected, GTC = Generalised Tonic Clonic Seizures

Supplementary Table 4. Detailed clinical features of individuals with *NARS1* Homozygous mutation c.1633C>T, p.R545C including families 9-15.															
Family	9	9	10	11	11	11	11	12	12	13	13	14	15	15	15
Individual	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Origin	North India	North India	North India	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan	Pakistan
Consanguinity	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Gender / Age at follow-up	M / 33	M / 17	M / 8	F / 17	F / 17	M / 19	F / 23	M / 8	M /21	M / 6m	M / 6.5	M / 16	F / 30	M / 16	F / 13
Occipital Frontal Circumference (OFC)															
Birth	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Follow –up Percentile / SD	NA	NA	46.5cm <1p *(-4.2 SD)*	52cm 2p *(-2.2 SD)*	49.5cm <1p *(-4.5 SD)*	52cm 2p *(-2.1 SD)*	50.3cm <1p *(-3.8 SD)*	46cm <1p *(4.6SD)*	51cm <1p *(2.8SD)*	NA	49cm 2p *(-2.0SD)*	NA	50.8cm <1p *(-3.3SD)*	48.2cm <1p *(-4.5SD)*	48cm <1p *(-4.4SD)*
Microcephaly	Y	Y	Y	Y	Y	y	Y	Y	Y	Y	Y	N	Y	Y	Y
Developmental Milestones & Intellectual Disabilities															
GDD	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sitting (Months)	12m	12m	12m	12m	14m	12m	12m	NA	NA	NA	NA	7m	4y	1y	2y
Walking (Years)	NA	NA	2y	1y 8m	2y 8m	2y 8m	2y 8m	6y 6m	3y	3y	3y	2y	10y	2y	3y
Language	Delayed	Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed
Neurological Findings															
Intellectual Disability	Severe	Moderate	Severe	Severe	Severe	Severe	Profound	Severe	Severe	Severe	Severe	Moderate	Severe	Moderate	Moderate

Seizures	GTC	N	N	N	N	GTC	GTC	N	N	GTC	GTC	N	Partial	Partial	GTC
Ataxia	Y	Y	Y	Y	N	N	N	NA	NA	NA	NA	NA	Y	NA	NA
Peripheral Neuropathy	Y	Y	N	N	N	N	N	NA	NA	NA	NA	Y	NA	NA	NA
Imaging	NAD	NAD	NAD	NA	NA	NA	NA	NA	NA	NA	NA	NAD	NA	NA	NA
Clinical Findings															
Dysmorphic Features	Contractures	Scoliosis	Clinodactyly Syndactyly	Clinodactyly	NAD	Clinodactyly Short Limbs	NAD	NAD	NAD	NAD	NAD	NAD	Dysplastic ears Syndactyly	NA	NA
Tone	Reduced	Reduced	NAD	NAD	NAD	NAD	NAD	Reduced	Reduced	NA	NA	N	N	NA	NA
Power	Reduced	Reduced	NAD	NAD	Reduced	Reduced	Reduced	Reduced	Reduced	Reduced	Reduced	Reduced	Reduced	NA	NA
Sensation	Reduced	Reduced	NAD	NAD	NAD	NAD	NAD	NA	NA	NA	NA	Reduced	NA	NA	NA
Co-ordination	Ataxic	Ataxic	Ataxic	Ataxic	NAD	NAD	NAD	NA	NA	NA	NA	NA	Ataxic	NA	NA
Reflexes	NAD	NAD	NAD	Reduced	Reduced	NAD	Reduced	NA	NA	Reduced	Reduced	Reduced	Reduced	NA	NA

Y = Yes, N= No, M= Male, F= Female, NA= Not Available, p = percentile, SD = standard deviation, NAD = No abnormalities detected, GTC = Generalised Tonic Clonic Seizures

Supplementary Table 5. Detailed clinical features of individuals with *NARS1* mutations. Autosomal recessive inheritance. Families 16-20.									
Family	16	16	17	18	18	19	19	20	21
Individual	24	25	26	27	28	29	30	31	32
Inheritance	Homozygous	Homozygous	Homozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous
Variant	c.32G>C p. Arg11Pro	c.32G>C p. Arg11Pro	c.50C>T p.Thr17Met	c.1049T>C, p.Leu350Pro c.1264G>A, p.Ala422Thr	c.1049T>C, p.Leu350Pro c.1264G>A, p.Ala422Thr	c.1067A>C, p.Asp356Ala c.203dupA, p.Met69Aspfs*	c.1067A>C, p.Asp356Ala c.203dupA, p.Met69Aspfs*4	c.268C>T, p.Arg90* c.394G>T, p.Gly132Cys	c.1376 C>T, p.Thr459Ile c.178 A>G, p.Lys60Glu
Origin	Kosovo	Kosovo	Libya	German	German	Turkey	Turkey	Canada	USA
Consanguinity	N	N	Y	N	N	N	N	N	N
Gender / Age at follow-up	M / 2	F / 2y 6m	F/ 7	F / 15	F / 21	M/ 14y 2m	F / 7y 10m	F / 8	M/3
Occipital Frontal Circumference (OFC)									
Birth	34cm 20p	NA	NA	33cm 11p	NA	NA	NA	NA	NA
Follow –up Percentile / SD	43cm <1p *(-4SD)*	40.5cm <1p*(-5 SD)*	42cm <1p *(-7.7SD)*	NA	NA	49.2Cm <1p*(-3.4SD)*	47.5cm <1p*(-3.4SD)*	41 cm <1p *(-8.7SD)*	43.5cm <1p*(-5SD)*
Microcephaly	Y	Y	Y	Y	Y	Y	Y	Y	Y
Developmental Milestones & Intellectual Disabilities									
GDD	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sitting (Months)	N	N	NA	NA	NA	NA	NA	NA	10m

Walking (Years)	N	N	NA	3	3	3	2	6y 10m	3y
Language	N	N	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed	Severely Delayed
Neurological Findings									
ID	Profound	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe
Seizures	Myoclonic / GTC	GTC	GTC	GTC	N	GTC	GTC	GTC	GTC
Peripheral Neuropathy	NA	NA	NA	Y	Y	NA	NA	NA	NA
Ataxia	NA	NA	NA	Y	Y	NA	NA	Y	Y
Imaging	Delayed Myelination	Delayed Myelination	NA	NAD	NA	Thickening of gyri	NA	Thin corpus callosum Decreased white matter	Arachnoid Cyst
Clinical Features									
Dysmorphic Features	NAD	NAD	NA	NAD	NAD	NAD	NAD	Hypotelorism	NAD
Tone	Increased	NAD	NA	Reduced	Reduced	NA	NA	Reduced	Increased
Power	Reduced	Reduced	NA	Reduced	Reduced	NA	NA	N	N
Sensation	NA	NA	NA	Reduced	Reduced	NA	NA	N	N
Co-ordination	NA	NA	NA	Ataxic gait	Ataxic gait	NA	NA	Ataxic	Ataxic
Reflexes	Reduced	Reduced	NA	Reduced	NA	NA	NA	N	Increased
Other	N	N	N	N	N	N	N	Hip dysplasia	N

Y = Yes, N= No, M= Male, F= Female, NA= Not Available, p = percentile, SD = standard deviation, NAD = No abnormalities detected, ID = Intellectual Disability, GTC = Generalised Tonic Clonic Seizures, ASD = Autistic Spectrum Disorder

Supplementary Table 6. Primer sequences used in the *NARS1* chapter (4.4).

Nrs1-PJR-F	TTTGTTAAATCATACCTCGAGATGGCGGGATTGGAATCAAAAGTTT
Nrs1-PJR-R	GCCTCGCGAGTCGACCTCGAGTTAAGGTGTGCAACGTTCAGTAAATCG
Nrs1DelFw	CTCTAACGAGACTATAAGTTATCCAAGGCCGGTTATTTGATATTTAACATTTTCACTAAC TTCAAACTGTCCTTTTAAAACGGATCCCCGGGTTAATTAA
Nrs1DelRv	AAATTCTAAGTAAACAACATAGTTCGCCCACTGTTCAAACATTAAAAGCTACCCATTTCT TCGATATGGATAAACTTTGCGAATTCGAGCTCGTTTAAAC
Nrs1ck-L	ACTAGCCGAAATTTTGGAATCA
Nrs1ck-R	CTAACTGACTCGCACCTAGCCT
KanR	CGGATGTGATGTGAGAACTGTATCCTAGC
KanF	CGCTATACTGCTGTCGATTCG
NARS1_F	GCGTTAGAAGGATATAGAGGCCA
NARS1_R	ACCATCTCGCAACACCAGAAA
GAPDH_F	TGTGGGCATCAATGGATTTGG
GAPDH_R	ACACCATGTATTCCGGGTCAAT

Supplementary Table 7. Missense mutations in *NARS1* with the description of protein affects. The substrate in the figures is the aminoacyl-adenylate of Brugia malayi Asparaginyl-tRNA synthetase complexed with ATP: Mg and L-Asp-beta-NOH adenylate: PPi: Mg. The strongest effect in terms of catalysis would be expected for Arg322Leu and Asp356Ala.

Residue	**Predicted effect of the mutation on Asn*RS1* protein**
Arg11Pro	Unique domain of *AsnRS1* (UNE-N)
Thr17Met	Unique domain of *AsnRS1* (UNE-N)
Arg322Leu	Disturbance of the stabilization of aa-adenylate
Leu350Pro	Probably little effect on enzymatic activity and tRNA recognition, predictable effect on *AsnRS1* dimer formation
Asp356Ala	Mutant at the interface between the two enzymes, loss of the side chain of Asp356 and therefore probably an interaction with the dimmer becoming more unstable
Ala422Thr	Affects the 3 'end of the acceptor arm
Gly509Ser	Disrupts the end of the interface between the two subunits of *AsnRS1*
Arg545Cys	Potential disruption of the interaction with the acceptor arm of the tRNA

Supplementary Table 8. Mutagenesis primer sequences used in the *TARS1* chapter (4.5).

R131H_F	TGTGGGACCTGGACCACCCTCTGGAAGAAGA
R131H_R	TCTTCTTCCAGAGGGTGGTCCAGGTCCCACA
V372I_F	AGAGGATTCCAGGAGATAGTCACCCCAAACA
V372I_R	TGTTTGGGGTGACTATCTCCTGGAATCCTCT
R663Q_F	TGAATAAAAAGATTCAAAATGCACAGTTAGC
R663Q_R	GCTAACTGTGCATTTTGAATCTTTTTATTCA
Q639P_F	CCCAAAAGGTACGAC**C**ACAATTCCACGATGC
Q639P_R	GCATCGTGGAATTGTGGTCGTACCTTTTGGG
R619C_F	TTTTGGCTGTCCCCT**T**GCCAGGTAATGGTAG
R619C_R	CTACCATTACCTGGCAAGGGGACAGCCAAAA
D115N_F	AGTCAAGGCCTGGCCAACAACACCGTTATTG
D115N_R	CAATAACGGTGTTGTTGGCCAGGCCTTGACT
K363R_F	ATGAATATGCCCAAAGGGTACGACAACAATTC
K363R_R	GAATTGTTGTCGTACCCTTTGGGCATATTCAT

Supplementary Table 9. Primer sequences used in the *PIGS1* chapter (5.2).

Forward_XhoI RS	aattctcgagAGAGATTCCGATTGGGTGAG
Reverse_*Bam*HI RS	attggatccCCTTTCTTTCCTCCCAGATG
SD6 forward	TCTGAGTCACCTGGACAACC
SA2 reverse	ATCTCAGTGGTATTTGTGAGC

Supplementary Table 9. Clinical data of patients with biallelic pathogenic variants in *PIGS* reported in the literature a

	This study						Nguyen et al.,			
Paitient	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Individual 1a	Individual 1b	Individual 2a	Individual 2b
Family	1	2	3	4	4	5	Family 1	Family 1	Family 2	Family 2
Variant (NM_0033198.3)	c.174G>C; p.(Gln58His)	c.174G>C; p.(Gln58His)	c.1070G>A; p.(Gly357Asp)	c.986C>G; p.(Pro329Arg)	c.986C>G; p.(Pro329Arg)	c.1141_1164dup24; c.734G>A p.Trp245*	c.108G>A; p.(Trp36*), c.101T>C; p.(Leu34Pro)	c.108G>A; p.(Trp36*), c.101T>C; p.(Leu34Pro)	c.1316_1352delCCACCA CCCTTACCTCCCTGGCGC	c.1316_1352delCCACCA CCCTTACCTCCCTGGCGCA
inheritance/phase	homozygous	homozygous	homozygous	homozygous	homozygous	compound het	compound het	compound het	homozygous	homozygous
Sex	male	male	male	male	Female	female	male	male	male	male
Ethnicity	Pakistan	Pakistan	Pakistan	Egypt	Egypt	Chinese	European	European	Mexican	Mexican
Consanguinity	Yes	Yes	Yes	Yes	Yes	No	No	No	Likely^	Likely^
Age at last investigation	3y	1.5 y	1.5 y	8 y	9 m	1.8 y	5.6 y	5.6 y	7.5 y	9 mo
Premature death	No	Yes, 1.9 y	Yes, 2 y	No	Yes, 1 y	No	No	No	No	No
Prenatal and birth history	Preterm birth	Regular	Regular	Regular	Regular	Failure to progress	Twin-twin transfusion and jaundice	Twin-twin transfusion and jaundice	Fetal distress due to failed induction	Decreased fetal activity polyhydramnios,
DD/ID	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Microcephaly	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Congenital anomalies	No	No	No	No	No	Yes, assymetric kidneys	No	No	Umbilical and inguinal hernias, bilateral	No
Epilepsy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Seizure onset	1 mo	1 mo	2 mo	2 mo	1 mo	1 mo	8 months	1 year	NA	8 months
Seizure type	Multifocal clonic and myoclonic jerks	Multifocal clonic seizures	Focal/multifocal clonic seizures	Focal seizures	Focal seizures	Multifocal seizures	Febrile seizures	Febrile seizures	NA	NA
Seizure evolution	Myoclonic epilepsy	Tonic clonic	Tonic spasms, GTCS	GTCS	GTCS, myoclonic	Multifocal seizures	Lennox–Gastaux	Nonfebrile seizures	NA	NA
Seizure frequency	3-4x/day reduced to 1-2x/month	Multiple/day	3-4x/day increased to multiple/day	1x/month	1x/day	10x/day	NA	NA	NA	NA
EEG features	Multifocal epileptifocm activity, burst suppression	Multifocal generalized discharges, burst suppression	multifocal epileptiform discharges with slow background	diffusely slowed, low-voltage, monomorphic background activity	NA	multifocal slow-waves and sharp waves complexes more prominent over the right hemisphere on a globally slowed and	NA	NA	NA	NA
Status epilepticus	No	Yes	Yes	Yes, 4 times per year	Yes, multiple per month	Yes, multiple	NA	NA	NA	NA
Medications employed	Levetiracetam, nitrazepam, topiramate	Levetiracetam, topiramate, clonazepam	Levetiracetam, carbamazepine, valproate	Valproate, leviteracetum, topiramate,	Valproate, leviteracetum, topiramate,	Clobazam, Levetiracetam, Perampanel	Levetiracetam, clobazam, rufinamide	NA	NA	Valproate
Response to treatment	Refractory	Refractory	Refractory	Refractory	Refractory	Refractory	Partial	Partial	Refractory	Refractory
Seizure-free	No	No	No	No	No	No	No	No	Yes	NA
B6 trial and results	Yes, decreased seizure frequency	No	No	No	No	Yes, improved development	No	No	Yes, seizure-free	Yes, NA
Behavioral anomalies	Mouth twitching	No	No	ASD,	Excessive crying	No	NA	NA	No	No
Hypotonia	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other neurologic features	No	No	Appendicular	Hyperreflexia	Rigidity	No	Ataxia	Ataxia	Ataxia, hyporeflexia	Ataxia, hyporeflexia
Coarse facial features	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hearing loss	Yes	No	No	No	No	No	Yes	Yes	No	NA
Impaired vision	Yes	Yes	No	No	No	Yes, cortical blindness	Yes, cortical blindness	Yes, cortical blindness	Yes	No
Other ophthalmologic	No	No	No	No	No	Astigmatism	Nystagmus	Nystagmus	Nystagmus	No
Gastrointestinal problems	No	No	No	No	Dysphagia	No	No	No	Dysphagia, hepatomegaly	Gastrointestinal reflux, hepatomegaly
Musculoskeletal abnormalities	No	No	No	Progressive knees contractures	No	No	brachydactyly, clinodactyly, short fourth metacarpals	brachydactyly, clinodactyly, short fourth metacarpals	Pectus carinatum, clinodactyly, short phalanges, scoliosis,	Pectus carniatum, hips hyperlaxity, clinodactyly
Cerebral atrophy	Yes, fronto-temporal	Yes, frontal	Yes, diffuse	Yes, ++	Yes, diffuse	Yes, frontotemporal	No	No	Yes, diffuse cortical	Yes, diffuse cortical
Cerebellar atrophy	No	Yes, superior	No	No	No	Yes, vermis	Yes	Yes	No	No
Other MRI findings	Frontal dysgyria	No	hypoplasia of the corpus callosum and	Delayed myelination	Subdural hygromas	Pontine hypoplasia	No	No	No	No

^parents from the same small rural community

ASD = autism spectrum disorder; GTCS = generalized tonic-clonic seizures; mo = months; NA = not available;

Supplementary Table 10. Clinical data of patients with mono-allelic and biallelic pathogenic variants in *SLITRK3* reported in our cohort.

	Family I			Family II	Family III
SLITRK3 variant (NM_014926.4)	c.1816G>T; p.(Glu606Ter)			c.1696T>C; p.(Cys566Arg)	c.660_684del25; p.(Gly221Trp*fs*X15)
inheritance/phase	Biallelic			De novo	De novo (mosaic) 11%of 107 sequenicng read from blood sample.
Current age, gender	7y, M	6y, M	5y, F	5y, M	16y, F
Ethnicity	Pakistani			Montenegro/Australian	African American
Consanguinity	+			No (but ROH on microarray)	No
History of miscarriages	+ (1)			1 miscarriage (first trimester)	No
Prenatal history	H/O HIE-I	insignificant	insignificant	Full-term, scheduled C-section	NA
Neonatal course	remain admitted for 2days	Normal	Normal	Normal	NA
Short stature	+ (-2 SDS)	+ (-2 SDS)	+ (-2.7 SDS)	No	+ (-5.9 SDS)
Microcephaly	-	-	+ (-2.9 SDS)	No	47cm (-5 SDS)
Dysmorphic features	-	-	-	-	larger ear, prominent nose, larger mouth, short philtrum, everted upper lip, camptodactyly in fingers, flat feet
Global DD/ID	+	+	+	-	+
Unassisted walk	+	+	+	+	Yes
Speech	Few words	Few words	Absent	Appropriate for age	Absent
Behavioral anomalies	-	-	-	-	breath holding, difficulty staying asleep, teeth grinding, hand rubbing, laughing fits, and crying spells
Seizures	+	+	+	+	+
Type, onset	GTCS, 6m	GTS and ES, 2y	GTC, 18m	FS, 3y	staring spells, 9-12m
Frequency	1-10/day	0-5/day	7-8/day	1/day	1/year, now seizure free for 1.5y
Duration	5 min	1-2 min	4-5 min	<1 min	1-2 min
Post-ictal status	+ +	+ +	-	+	+
Evolution	GTCS	GTS and ES	GTC	GTCS	GTCS
Status epilepticus	2 times	once	No	No	No
Trialled drugs	Pheno, VPA, LEV	VPA, LEV	LEV	Keppra 1.5ml BID	Keppra and Clonazepam
Response to treatment	Controlled	Controlled	Controlled	Controlled (last sz in 2018)	Controlled
EEG	Left parietal focal epileptogenic activity	Diffuse encephalopathy	Diffuse encephalopathy	Background slowing, multifocal IEDs	generalized atypical spike slow waves: (1) Frequent bursts of 1.5 - 2 Hz generalized, irregular slow spike-and-slow wave complexes. (2) Occasional bursts of generalized paroxysmal fast activity. (3) Intermittent generalized delta activity. (4) Lack of a well-sustained/modulated PDR, lack of well-organized APG and moderate background slowing.
Hypotonia	-	-	-	+, mild	+
Spasticity	-	+	-	-	+
Psychomotor regression	-	-	-	-	yes, knew a few words and signs by 12m but lost all around 12m
Opthalmological findings	RPE abnormalities	-	-	-	cortical vision impairment
MRI anomalies	-	NA	NA	+	1. Mild ventricular enlargement of prominent sulci suggestive of low brain volumes. 2. Thinned corpus callosum. 3. There is evidence of subependymal nodular heterotopia along the bilateral lateral ventricles anteriorly 4. Irregular border to the ventricular system, raising the possibility of PVL.
White matter alterations	-	NA	NA	+ (subcortical T2-hyperintensities)	volume loss in the periventricular white matter
Other	-	NA	NA	Optic nerve head flattening	NA
Metabolic abnormalities	↑lactate and ammonia	NA	NA	NA	NA

Abbreviations: ES (extensor spasms); F (female); FS (febrile seizures), GTCS (generalized tonic-clonic seizures); IED (interictal epileptiform discharge); M (male); NA (not available); RPE (retinal pigmented epithelium); y (years)

References

1. Efthymiou, S., Salpietro, V., Malintan, N., Poncelet, M., Kriouile, Y., Fortuna, S., De Zorzi, R., Payne, K., Henderson, L.B., Cortese, A., et al. (2019). Biallelic mutations in neurofascin cause neurodevelopmental impairment and peripheral demyelination. Brain 142, 2948-2964.
2. Smigiel, R., Sherman, D.L., Rydzanicz, M., Walczak, A., Mikolajkow, D., Krolak-Olejnik, B., Kosinska, J., Gasperowicz, P., Biernacka, A., Stawinski, P., et al. (2018). Homozygous mutation in the Neurofascin gene affecting the glial isoform of Neurofascin causes severe neurodevelopment disorder with hypotonia, amimia and areflexia. Hum Mol Genet.
3. Anazi, S., Maddirevula, S., Faqeih, E., Alsedairy, H., Alzahrani, F., Shamseldin, H.E., Patel, N., Hashem, M., Ibrahim, N., Abdulwahab, F., et al. (2017). Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. Molecular psychiatry 22, 615-624.
4. Monfrini, E., Straniero, L., Bonato, S., Monzio Compagnoni, G., Bordoni, A., Dilena, R., Rinchetti, P., Silipigni, R., Ronchi, D., Corti, S., et al. (2019). Neurofascin (NFASC) gene mutation causes autosomal recessive ataxia with demyelinating neuropathy. Parkinsonism & related disorders.
5. Manole, A., Efthymiou, S., O'Connor, E., Mendes, M.I., Jennings, M., Maroofian, R., Davagnanam, I., Mankad, K., Lopez, M.R., Salpietro, V., et al. (2020). De Novo and Bi-allelic Pathogenic Variants in NARS1 Cause Neurodevelopmental Delay Due to Toxic Gain-of-Function and Partial Loss-of-Function Effects. Am J Hum Genet 107, 311-324.
6. Karakaya, M., and Wirth, B. (2019). Hereditary nodo-paranodopathies: genomic variants, not just autoantibodies, hit the protein. Brain 142, 2895-2897.
7. Bucelli, R.C., Arhzaouy, K., Pestronk, A., Pittman, S.K., Rojas, L., Sue, C.M., Evila, A., Hackman, P., Udd, B., Harms, M.B., et al. (2015). SQSTM1 splice site mutation in distal myopathy with rimmed vacuoles. Neurology 85, 665-674.
8. Cortese, A., Machado, P., Morrow, J., Dewar, L., Hiscock, A., Miller, A., Brady, S., Hilton-Jones, D., Parton, M., and Hanna, M.G. (2013). Longitudinal observational study of sporadic inclusion body myositis: implications for clinical trials. Neuromuscular disorders : NMD 23, 404-412.
9. Evila, A., Arumilli, M., Udd, B., and Hackman, P. (2016). Targeted next-generation sequencing assay for detection of mutations in primary myopathies. Neuromuscular disorders : NMD 26, 7-15.
10. Hermans, M.C., Faber, C.G., Bekkers, S.C., de Die-Smulders, C.E., Gerrits, M.M., Merkies, I.S., Snoep, G., Pinto, Y.M., and Schalla, S. (2012). Structural and functional cardiac changes in myotonic dystrophy type 1: a cardiovascular magnetic resonance study. Journal of Cardiovascular Magnetic Resonance 14, 48.
11. Hunter, J.M., Ahearn, M.E., Balak, C.D., Liang, W.S., Kurdoglu, A., Corneveaux, J.J., Russell, M., Huentelman, M.J., Craig, D.W., Carpten, J., et al. (2015). Novel pathogenic variants and genes for myopathies identified by whole exome sequencing. Molecular genetics & genomic medicine 3, 283-301.
12. Izumi, R., Warita, H., Niihori, T., Takahashi, T., Tateyama, M., Suzuki, N., Nishiyama, A., Shirota, M., Funayama, R., Nakayama, K., et al. (2015). Isolated inclusion body myopathy caused by a multisystem proteinopathy-linked hnRNPA1 mutation. Neurology Genetics 1, e23.
13. Murphy, S.M., Laura, M., Fawcett, K., Pandraud, A., Liu, Y.T., Davidson, G.L., Rossor, A.M., Polke, J.M., Castleman, V., Manji, H., et al. (2012). Charcot-Marie-Tooth disease: frequency of genetic subtypes and guidelines for genetic testing. Journal of neurology, neurosurgery, and psychiatry 83, 706-710.
14. Rossor, A.M., Evans, M.R., and Reilly, M.M. (2015). A practical approach to the genetic neuropathies. Practical neurology 15, 187-198.
15. Blokhuis, A.M., Groen, E.J., Koppers, M., van den Berg, L.H., and Pasterkamp, R.J. (2013). Protein aggregation in amyotrophic lateral sclerosis. Acta neuropathologica 125, 777-794.
16. Borghero, G., Pugliatti, M., Marrosu, F., Marrosu, M.G., Murru, M.R., Floris, G., Cannas, A., Occhineri, P., Cau, T.B., Loi, D., et al. (2016). TBK1 is associated with ALS and ALS-FTD in Sardinian patients. Neurobiology of aging.
17. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245-256.
18. Gitler, A.D., and Shorter, J. (2011). RNA-binding proteins with prion-like domains in ALS and FTLD-U. Prion 5, 179-187.
19. Lim, L., Wei, Y., Lu, Y., and Song, J. (2016). ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. PLoS biology 14, e1002338.
20. Castro-Fernandez, C., Arias, M., Blanco-Arias, P., Santome-Collazo, L., Amigo, J., Carracedo, A., and Sobrido, M.J. (2015). Targeted NGS meets expert clinical characterization: Efficient diagnosis of spastic paraplegia type 11. Applied & translational genomics 5, 33-36.
21. Lynch, D.S., Koutsis, G., Tucci, A., Panas, M., Baklou, M., Breza, M., Karadima, G., and Houlden, H. (2015). Hereditary spastic paraplegia in Greece: characterisation of a previously unexplored population using next-generation sequencing. European journal of human genetics : EJHG.
22. Rinaldi, C., Schmidt, T., Situ, A.J., Johnson, J.O., Lee, P.R., Chen, K.L., Bott, L.C., Fado, R., Harmison, G.H., Parodi, S., et al. (2015). Mutation in CPT1C Associated With Pure Autosomal Dominant Spastic Paraplegia. JAMA neurology 72, 561-570.
23. Salinas, S., Proukakis, C., Crosby, A., and Warner, T.T. (2008). Hereditary spastic paraplegia: clinical features and pathogenetic mechanisms. The Lancet Neurology 7, 1127-1138.
24. Stevanin, G., Santorelli, F.M., Azzedine, H., Coutinho, P., Chomilier, J., Denora, P.S., Martin, E., Ouvrard-Hernandez, A.M., Tessa, A., Bouslam, N., et al. (2007). Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. Nature genetics 39, 366-372.
25. Filla, A., and De Michele, G. (2012). Overview of autosomal recessive ataxias. Handbook of clinical neurology 103, 265-274.

26. Maksemous, N., Roy, B., Smith, R.A., and Griffiths, L.R. (2016). Next-generation sequencing identifies novel CACNA1A gene mutations in episodic ataxia type 2. Molecular genetics & genomic medicine 4, 211-222.
27. Synofzik, M., Smets, K., Mallaret, M., Di Bella, D., Gallenmuller, C., Baets, J., Schulze, M., Magri, S., Sarto, E., Mustafa, M., et al. (2016). SYNE1 ataxia is a common recessive ataxia with major non-cerebellar features: a large scale multi-centre study. Brain.
28. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic acids research 44, D457-462.
29. Chae, J.H., Vasta, V., Cho, A., Lim, B.C., Zhang, Q., Eun, S.H., and Hahn, S.H. (2015). Utility of next generation sequencing in genetic diagnosis of early onset neuromuscular disorders. Journal of medical genetics 52, 208-216.
30. Lindquist, S.G., Moller, L.B., Dali, C.I., Marner, L., Kamsteeg, E.J., Nielsen, J.E., and Hjermind, L.E. (2016). A Novel TTBK2 De Novo Mutation in a Danish Family with Early-Onset Spinocerebellar Ataxia. Cerebellum.
31. Siekierska, A., Isrie, M., Liu, Y., Scheldeman, C., Vanthillo, N., Lagae, L., de Witte, P.A., Van Esch, H., Goldfarb, M., and Buyse, G.M. (2016). Gain-of-function FHF1 mutation causes early-onset epileptic encephalopathy with cerebellar atrophy. Neurology.
32. Tzoulis, C., Sztromwasser, P., Johansson, S., Gjerde, I.O., Knappskog, P., and Bindoff, L.A. (2016). PNKP Mutations Identified by Whole-Exome Sequencing in a Norwegian Patient with Sporadic Ataxia and Edema. Cerebellum.
33. van de Warrenburg, B.P., Schouten, M.I., de Bot, S.T., Vermeer, S., Meijer, R., Pennings, M., Gilissen, C., Willemsen, M.A., Scheffer, H., and Kamsteeg, E.J. (2016). Clinical exome sequencing for cerebellar ataxia and spastic paraplegia uncovers novel gene-disease associations and unanticipated rare disorders. European journal of human genetics : EJHG.
34. Wang, W., Wang, C., Dawson, D.B., Thorland, E.C., Lundquist, P.A., Eckloff, B.W., Wu, Y., Baheti, S., Evans, J.M., Scherer, S.S., et al. (2016). Target-enrichment sequencing and copy number evaluation in inherited polyneuropathy. Neurology 86, 1762-1771.
35. Sanger, F., Nicklen, S., and Coulson, A.R. (1992). DNA sequencing with chain-terminating inhibitors. 1977. Biotechnology (Reading, Mass) 24, 104-108.
36. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. Nature 508, 469-476.
37. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., et al. (2009). Finding the missing heritability of complex diseases. Nature 461, 747-753.
38. von Bubnoff, A. (2008). Next-generation sequencing: the race is on. Cell 132, 721-723.
39. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. The New England journal of medicine 369, 1502-1511.
40. Legati, A., Reyes, A., Nasca, A., Invernizzi, F., Lamantea, E., Tiranti, V., Garavaglia, B., Lamperti, C., Ardissone, A., Moroni, I., et al. (2016). New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. Biochimica et biophysica acta.
41. Endo, Y., Dong, M., Noguchi, S., Ogawa, M., Hayashi, Y.K., Kuru, S., Sugiyama, K., Nagai, S., Ozasa, S., Nonaka, I., et al. (2015). Milder forms of muscular dystrophy associated with POMGNT2 mutations. Neurology Genetics 1, e33.
42. Vasta, V., Merritt, J.L., 2nd, Saneto, R.P., and Hahn, S.H. (2012). Next-generation sequencing for mitochondrial diseases: a wide diagnostic spectrum. Pediatrics international : official journal of the Japan Pediatric Society 54, 585-601.
43. Legati, A., Reyes, A., Nasca, A., Invernizzi, F., Lamantea, E., Tiranti, V., Garavaglia, B., Lamperti, C., Ardissone, A., Moroni, I., et al. (2016). New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. Biochimica et biophysica acta 1857, 1326-1335.
44. Alazami, A.M., Patel, N., Shamseldin, H.E., Anazi, S., Al-Dosari, M.S., Alzahrani, F., Hijazi, H., Alshammari, M., Aldahmesh, M.A., Salih, M.A., et al. (2015). Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. Cell reports 10, 148-161.
45. Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., et al. (2010). Exome sequencing identifies the cause of a mendelian disorder. Nature genetics 42, 30-35.
46. Kaplan, J.C., and Hamroun, D. (2014). The 2015 version of the gene table of monogenic neuromuscular disorders (nuclear genome). Neuromuscular disorders : NMD 24, 1123-1153.
47. Liu, Y.T., Hersheson, J., Plagnol, V., Fawcett, K., Duberley, K.E., Preza, E., Hargreaves, I.P., Chalasani, A., Laura, M., Wood, N.W., et al. (2014). Autosomal-recessive cerebellar ataxia caused by a novel ADCK3 mutation that elongates the protein: clinical, genetic and biochemical characterisation. Journal of neurology, neurosurgery, and psychiatry 85, 493-498.
48. Gang, Q., Bettencourt, C., Machado, P.M., Fox, Z., Brady, S., Healy, E., Parton, M., Holton, J.L., Hilton-Jones, D., Shieh, P.B., et al. (2015). The effects of an intronic polymorphism in TOMM40 and APOE genotypes in sporadic inclusion body myositis. Neurobiology of aging 36, 1766.e1761-1763.
49. Guerreiro, R., Bras, J., Hardy, J., and Singleton, A. (2014). Next generation sequencing techniques in neurological diseases: redefining clinical and molecular associations. Hum Mol Genet 23, R47-53.
50. Pasaniuc, B., Rohland, N., McLaren, P.J., Garimella, K., Zaitlen, N., Li, H., Gupta, N., Neale, B.M., Daly, M.J., Sklar, P., et al. (2012). Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nature genetics 44, 631-635.
51. Dixon-Salazar, T.J., Silhavy, J.L., Udpa, N., Schroth, J., Bielas, S., Schaffer, A.E., Olvera, J., Bafna, V., Zaki, M.S., Abdel-Salam, G.H., et al. (2012). Exome sequencing can improve diagnosis and alter patient management. Science translational medicine 4, 138ra178.
52. Botstein, D., and Risch, N. (2003). Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature genetics 33 Suppl, 228-237.

53. Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., et al. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome biology 10, R32.
54. Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. Nature reviews Genetics 12, 745-755.
55. Bell, C.J., Dinwiddie, D.L., Miller, N.A., Hateley, S.L., Ganusova, E.E., Mudge, J., Langley, R.J., Zhang, L., Lee, C.C., Schilkey, F.D., et al. (2011). Carrier testing for severe childhood recessive diseases by next-generation sequencing. Science translational medicine 3, 65ra64.
56. Goldstein, D.B., Allen, A., Keebler, J., Margulies, E.H., Petrou, S., Petrovski, S., and Sunyaev, S. (2013). Sequencing studies in human genetics: design and interpretation. Nature reviews Genetics 14, 460-470.
57. Kiezun, A., Garimella, K., Do, R., Stitziel, N.O., Neale, B.M., McLaren, P.J., Gupta, N., Sklar, P., Sullivan, P.F., Moran, J.L., et al. (2012). Exome sequencing and the genetic basis of complex traits. Nature genetics 44, 623-630.
58. Schuster, S.C. (2008). Next-generation sequencing transforms today's biology. Nature methods 5, 16-18.
59. Houlden, H., King, R.H., Hashemi-Nejad, A., Wood, N.W., Mathias, C.J., Reilly, M., and Thomas, P.K. (2001). A novel TRK A (NTRK1) mutation associated with hereditary sensory and autonomic neuropathy type V. Ann Neurol 49, 521-525.
60. Sailer, A., Scholz, S.W., Gibbs, J.R., Tucci, A., Johnson, J.O., Wood, N.W., Plagnol, V., Hummerich, H., Ding, J., Hernandez, D., et al. (2012). Exome sequencing in an SCA14 family demonstrates its utility in diagnosing heterogeneous diseases. Neurology 79, 127-131.
61. Singleton, A.B., Hardy, J., Traynor, B.J., and Houlden, H. (2010). Towards a complete resolution of the genetic architecture of disease. Trends Genet 26, 438-442.
62. Hernandez, D.G., Reed, X., and Singleton, A.B. (2016). Genetics in Parkinson disease: Mendelian versus non-Mendelian inheritance. Journal of neurochemistry 139 Suppl 1, 59-74.
63. Levy, S.E., and Myers, R.M. (2016). Advancements in Next-Generation Sequencing. Annual review of genomics and human genetics 17, 95-115.
64. Zollo, M., Ahmed, M., Ferrucci, V., Salpietro, V., Asadzadeh, F., Carotenuto, M., Maroofian, R., Al-Amri, A., Singh, R., Scognamiglio, I., et al. (2017). PRUNE is crucial for normal brain development and mutated in microcephaly with neurodevelopmental impairment. Brain 140, 940-952.
65. Salpietro, V., Perez-Duenas, B., Nakashima, K., San Antonio-Arce, V., Manole, A., Efthymiou, S., Vandrovcova, J., Bettencourt, C., Mencacci, N.E., Klein, C., et al. (2018). A homozygous loss-of-function mutation in PDE2A associated to early-onset hereditary chorea. Mov Disord 33, 482-488.
66. Chelban, V., Patel, N., Vandrovcova, J., Zanetti, M.N., Lynch, D.S., Ryten, M., Botia, J.A., Bello, O., Tribollet, E., Efthymiou, S., et al. (2017). Mutations in NKX6-2 Cause Progressive Spastic Ataxia and Hypomyelination. Am J Hum Genet 100, 969-977.
67. Hoskinson, D.C., Dubuc, A.M., and Mason-Suares, H. (2017). The current state of clinical interpretation of sequence variants. Current opinion in genetics & development 42, 33-39.
68. Rehm, H.L., Bale, S.J., Bayrak-Toydemir, P., Berg, J.S., Brown, K.K., Deignan, J.L., Friez, M.J., Funke, B.H., Hegde, M.R., and Lyon, E. (2013). ACMG clinical laboratory standards for next-generation sequencing. Genetics in medicine : official journal of the American College of Medical Genetics 15, 733-747.
69. Efthymiou, S., Manole, A., and Houlden, H. (2016). Next-generation sequencing in neuromuscular diseases. Current opinion in neurology 29, 527-536.
70. Maddalena, A., Bale, S., Das, S., Grody, W., and Richards, S. (2005). Technical standards and guidelines: molecular genetic testing for ultra-rare disorders. Genetics in medicine : official journal of the American College of Medical Genetics 7, 571-583.
71. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (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. Genetics in medicine : official journal of the American College of Medical Genetics 17, 405-424.
72. Devaux, J.J. (2012). Antibodies to gliomedin cause peripheral demyelinating neuropathy and the dismantling of the nodes of Ranvier. The American journal of pathology 181, 1402-1413.
73. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., and Shaw, D.E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78, 1950-1958.
74. Bussi, G., Donadio, D., and Parrinello, M. (2007). Canonical sampling through velocity rescaling. The Journal of chemical physics 126, 014101.
75. Hess, B. (2008). P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. Journal of chemical theory and computation 4, 116-122.
76. Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., van der Spoel, D., et al. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics (Oxford, England) 29, 845-854.
77. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science (New York, NY) 348, 648-660.
78. Carithers, L.J., Ardlie, K., Barcus, M., Branton, P.A., Britton, A., Buia, S.A., Compton, C.C., DeLuca, D.S., Peter-Demchok, J., Gelfand, E.T., et al. (2015). A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. Biopreservation and biobanking 13, 311-319.
79. Leek, J.T., and Storey, J.D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet 3, 1724-1735.
80. Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics (Oxford, England) 8, 118-127.

81. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics 9, 559.
82. Fabregat, A., Korninger, F., Viteri, G., Sidiropoulos, K., Marin-Garcia, P., Ping, P., Wu, G., Stein, L., D'Eustachio, P., and Hermjakob, H. (2018). Reactome graph database: Efficient access to complex pathway data. PLoS Comput Biol 14, e1005968.
83. Thomas-Jinu, S., Gordon, P.M., Fielding, T., Taylor, R., Smith, B.N., Snowden, V., Blanc, E., Vance, C., Topp, S., Wong, C.H., et al. (2017). Non-nuclear Pool of Splicing Factor SFPQ Regulates Axonal Transcripts Required for Normal Motor Development. Neuron 94, 322-336.e325.
84. Verma, P.K., and El-Harouni, A.A. (2015). Review of literature: genes related to postaxial polydactyly. Front Pediatr 3, 8.
85. Deng, H., Tan, T., and Yuan, L. (2015). Advances in the molecular genetics of non-syndromic polydactyly. Expert Rev Mol Med 17, e18.
86. Botia, J.A., Vandrovcova, J., Forabosco, P., Guelfi, S., D'Sa, K., United Kingdom Brain Expression, C., Hardy, J., Lewis, C.M., Ryten, M., and Weale, M.E. (2017). An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. BMC Syst Biol 11, 47.
87. Forabosco, P., Ramasamy, A., Trabzuni, D., Walker, R., Smith, C., Bras, J., Levine, A.P., Hardy, J., Pocock, J.M., Guerreiro, R., et al. (2013). Insights into TREM2 biology by network analysis of human brain gene expression data. Neurobiology of aging 34, 2699-2714.
88. Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic acids research 35, W193-200.
89. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498-2504.
90. Retterer, K., Juusola, J., Cho, M.T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K.G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. Genetics in medicine : official journal of the American College of Medical Genetics 18, 696-704.
91. Harripaul, R., Vasli, N., Mikhailov, A., Rafiq, M.A., Mittal, K., Windpassinger, C., Sheikh, T.I., Noor, A., Mahmood, H., Downey, S., et al. (2018). Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. Mol Psychiatry 23, 973-984.
92. Cooper, G.M., Stone, E.A., Asimenos, G., Program, N.C.S., Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome Res 15, 901-913.
93. Kuhn, R.M., Karolchik, D., Zweig, A.S., Wang, T., Smith, K.E., Rosenbloom, K.R., Rhead, B., Raney, B.J., Pohl, A., Pheasant, M., et al. (2009). The UCSC Genome Browser Database: update 2009. Nucleic acids research 37, D755-761.
94. Moreno, M.B., Duran, A., and Ribas, J.C. (2000). A family of multifunctional thiamine-repressible expression vectors for fission yeast. Yeast 16, 861-872.
95. Rodriguez-Lopez, M., Cotobal, C., Fernandez-Sanchez, O., Borbaran Bravo, N., Oktriani, R., Abendroth, H., Uka, D., Hoti, M., Wang, J., Zaratiegui, M., et al. (2017). A CRISPR/Cas9-based method and primer design tool for seamless genome editing in fission yeast [version 3; referees: 2 approved]. 1.
96. Bahler, J., Wu, J.Q., Longtine, M.S., Shah, N.G., McKenzie, A., 3rd, Steever, A.B., Wach, A., Philippsen, P., and Pringle, J.R. (1998). Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast 14, 943-951.
97. Sato, M., Dhut, S., and Toda, T. (2005). New drug-resistant cassettes for gene disruption and epitope tagging in Schizosaccharomyces pombe. Yeast 22, 583-591.
98. Meyer, J., Novak, M., Hamel, A., and Rosenberg, K. (2014). Extraction and analysis of cortisol from human and monkey hair. J Vis Exp, e50882.
99. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics (Oxford, England) 31, 166-169.
100. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15, 550.
101. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry 25, 1605-1612.
102. Flex, E., Niceta, M., Cecchetti, S., Thiffault, I., Au, M.G., Capuano, A., Piermarini, E., Ivanova, A.A., Francis, J.W., Chillemi, G., et al. (2016). Biallelic Mutations in TBCD, Encoding the Tubulin Folding Cofactor D, Perturb Microtubule Dynamics and Cause Early-Onset Encephalopathy. Am J Hum Genet 99, 962-973.
103. Bauer, C.K., Calligari, P., Radio, F.C., Caputo, V., Dentici, M.L., Falah, N., High, F., Pantaleoni, F., Barresi, S., Ciolfi, A., et al. (2018). Mutations in KCNK4 that Affect Gating Cause a Recognizable Neurodevelopmental Syndrome. Am J Hum Genet 103, 621-630.
104. Flex, E., Martinelli, S., Van Dijck, A., Ciolfi, A., Cecchetti, S., Coluzzi, E., Pannone, L., Andreoli, C., Radio, F.C., Pizzi, S., et al. (2019). Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. Am J Hum Genet 105, 493-508.
105. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43, 11 10 11-11 10 33.
106. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997v1 [q-bioGN].
107. Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80-92.

108. Liu, X., Wu, C., Li, C., and Boerwinkle, E. (2016). dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Human mutation 37, 235-241.
109. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nature genetics 46, 310-315.
110. Jagadeesh, K.A., Wenger, A.M., Berger, M.J., Guturu, H., Stenson, P.D., Cooper, D.N., Bernstein, J.A., and Bejerano, G. (2016). M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. Nature genetics 48, 1581-1586.
111. Li, Q., and Wang, K. (2017). InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet 100, 267-280.
112. Oprescu, S.N., Griffin, L.B., Beg, A.A., and Antonellis, A. (2017). Predicting the pathogenicity of aminoacyl-tRNA synthetase mutations. Methods 113, 139-151.
113. Chien, C.I., Chen, Y.W., Wu, Y.H., Chang, C.Y., Wang, T.L., and Wang, C.C. (2014). Functional substitution of a eukaryotic glycyl-tRNA synthetase with an evolutionarily unrelated bacterial cognate enzyme. PLoS One 9, e94659.
114. Boeke, J.D., LaCroute, F., and Fink, G.R. (1984). A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet 197, 345-346.
115. Mencacci, N.E., Kamsteeg, E.J., Nakashima, K., R'Bibo, L., Lynch, D.S., Balint, B., Willemsen, M.A., Adams, M.E., Wiethoff, S., Suzuki, K., et al. (2016). De Novo Mutations in PDE10A Cause Childhood-Onset Chorea with Bilateral Striatal Lesions. Am J Hum Genet 98, 763-771.
116. Makrythanasis, P., Maroofian, R., Stray-Pedersen, A., Musaev, D., Zaki, M.S., Mahmoud, I.G., Selim, L., Elbadawy, A., Jhangiani, S.N., Coban Akdemir, Z.H., et al. (2018). Biallelic variants in KIF14 cause intellectual disability with microcephaly. European journal of human genetics : EJHG 26, 330-339.
117. Vona, B., Mazaheri, N., Lin, S.-J., Dunbar, L.A., Maroofian, R., Azaiez, H., Booth, K.T., Vitry, S., Rad, A., Varshney, P., et al. (2020). Biallelic mutation of CLRN2 causes non-syndromic hearing loss in humans. bioRxiv, 2020.2007.2029.222828.
118. Tompson, S.W., and Young, T.L. (2017). Assaying the Effects of Splice Site Variants by Exon Trapping in a Mammalian Cell Line. Bio Protoc 7.
119. Doll, J., Kolb, S., Schnapp, L., Rad, A., Ruschendorf, F., Khan, I., Adli, A., Hasanzadeh, A., Liedtke, D., Knaup, S., et al. (2020). Novel Loss-of-Function Variants in CDC14A are Associated with Recessive Sensorineural Hearing Loss in Iranian and Pakistani Patients. International journal of molecular sciences 21.
120. Gu, X., Mao, X., Lussier, M.P., Hutchison, M.A., Zhou, L., Hamra, F.K., Roche, K.W., and Lu, W. (2016). GSG1L suppresses AMPA receptor-mediated synaptic transmission and uniquely modulates AMPA receptor kinetics in hippocampal neurons. Nat Commun 7, 10873.
121. Rossor, A.M., Tomaselli, P.J., and Reilly, M.M. (2016). Recent advances in the genetic neuropathies. Current opinion in neurology 29, 537-548.
122. Rossor, A.M., Carr, A.S., Devine, H., Chandrashekar, H., Pelayo-Negro, A.L., Pareyson, D., Shy, M.E., Scherer, S.S., and Reilly, M.M. (2017). Peripheral neuropathy in complex inherited diseases: an approach to diagnosis. Journal of neurology, neurosurgery, and psychiatry 88, 846-863.
123. Choi, M., Scholl, U.I., Ji, W., Liu, T., Tikhonova, I.R., Zumbo, P., Nayir, A., Bakkaloglu, A., Ozen, S., Sanjad, S., et al. (2009). Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 106, 19096-19101.
124. Montenegro, G., Powell, E., Huang, J., Speziani, F., Edwards, Y.J., Beecham, G., Hulme, W., Siskind, C., Vance, J., Shy, M., et al. (2011). Exome sequencing allows for rapid gene identification in a Charcot-Marie-Tooth family. Ann Neurol 69, 464-470.
125. Rossor, A.M., Polke, J.M., Houlden, H., and Reilly, M.M. (2013). Clinical implications of genetic advances in Charcot-Marie-Tooth disease. Nature reviews Neurology 9, 562-571.
126. Pipis, M., Rossor, A.M., Laura, M., and Reilly, M.M. (2019). Next-generation sequencing in Charcot-Marie-Tooth disease: opportunities and challenges. Nature reviews Neurology 15, 644-656.
127. Ghosh, A., Sherman, D.L., and Brophy, P.J. (2018). The Axonal Cytoskeleton and the Assembly of Nodes of Ranvier. The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry 24, 104-110.
128. Sherman, D.L., Tait, S., Melrose, S., Johnson, R., Zonta, B., Court, F.A., Macklin, W.B., Meek, S., Smith, A.J., Cottrell, D.F., et al. (2005). Neurofascins are required to establish axonal domains for saltatory conduction. Neuron 48, 737-742.
129. Taylor, A.M., Shi, Q., and Bhat, M.A. (2018). Simultaneous Ablation of Neuronal Neurofascin and Ankyrin G in Young and Adult Mice Reveals Age-Dependent Increase in Nodal Stability in Myelinated Axons and Differential Effects on the Lifespan. eNeuro 5.
130. Querol, L., Siles, A.M., Alba-Rovira, R., Jauregui, A., Devaux, J., Faivre-Sarrailh, C., Araque, J., Rojas-Garcia, R., Diaz-Manera, J., Cortes-Vicente, E., et al. (2017). Antibodies against peripheral nerve antigens in chronic inflammatory demyelinating polyradiculoneuropathy. Scientific reports 7, 14411.
131. Vallat, J.M., Mathis, S., Magy, L., Bounolleau, P., Skarzynski, M., Heitzmann, A., Manso, C., Devaux, J., and Uncini, A. (2018). Subacute nodopathy with conduction blocks and anti-neurofascin 140/186 antibodies: an ultrastructural study. Brain 141, e56.
132. Maluenda, J., Manso, C., Quevarec, L., Vivanti, A., Marguet, F., Gonzales, M., Guimiot, F., Petit, F., Toutain, A., Whalen, S., et al. (2016). Mutations in GLDN, Encoding Gliomedin, a Critical Component of the Nodes of Ranvier, Are Responsible for Lethal Arthrogryposis. Am J Hum Genet 99, 928-933.
133. Hengel, H., Magee, A., Mahanjah, M., Vallat, J.M., Ouvrier, R., Abu-Rashid, M., Mahamid, J., Schule, R., Schulze, M., Krageloh-Mann, I., et al. (2017). CNTNAP1 mutations cause CNS hypomyelination and neuropathy with or without arthrogryposis. Neurology Genetics 3, e144.

134. Buttermore, E.D., Piochon, C., Wallace, M.L., Philpot, B.D., Hansel, C., and Bhat, M.A. (2012). Pinceau organization in the cerebellum requires distinct functions of neurofascin in Purkinje and basket neurons during postnatal development. The Journal of neuroscience : the official journal of the Society for Neuroscience 32, 4724-4742.
135. Labasque, M., Hivert, B., Nogales-Gadea, G., Querol, L., Illa, I., and Faivre-Sarrailh, C. (2014). Specific contactin N-glycans are implicated in neurofascin binding and autoimmune targeting in peripheral neuropathies. The Journal of biological chemistry 289, 7907-7918.
136. Zhang, Y., Sloan, S.A., Clarke, L.E., Caneda, C., Plaza, C.A., Blumenthal, P.D., Vogel, H., Steinberg, G.K., Edwards, M.S., Li, G., et al. (2016). Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron 89, 37-53.
137. Zeisel, A., Hochgerner, H., Lonnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular Architecture of the Mouse Nervous System. Cell 174, 999-1014.e1022.
138. Laquerriere, A., Maluenda, J., Camus, A., Fontenas, L., Dieterich, K., Nolent, F., Zhou, J., Monnier, N., Latour, P., Gentil, D., et al. (2014). Mutations in CNTNAP1 and ADCY6 are responsible for severe arthrogryposis multiplex congenita with axoglial defects. Hum Mol Genet 23, 2279-2289.
139. Querol, L., Nogales-Gadea, G., Rojas-Garcia, R., Diaz-Manera, J., Pardo, J., Ortega-Moreno, A., Sedano, M.J., Gallardo, E., Berciano, J., Blesa, R., et al. (2014). Neurofascin IgG4 antibodies in CIDP associate with disabling tremor and poor response to IVIg. Neurology 82, 879-886.
140. Ogata, H., Yamasaki, R., Hiwatashi, A., Oka, N., Kawamura, N., Matsuse, D., Kuwahara, M., Suzuki, H., Kusunoki, S., Fujimoto, Y., et al. (2015). Characterization of IgG4 anti-neurofascin 155 antibody-positive polyneuropathy. Annals of clinical and translational neurology 2, 960-971.
141. Delmont, E., Manso, C., Querol, L., Cortese, A., Berardinelli, A., Lozza, A., Belghazi, M., Malissart, P., Labauge, P., Taieb, G., et al. (2017). Autoantibodies to nodal isoforms of neurofascin in chronic inflammatory demyelinating polyneuropathy. Brain 140, 1851-1858.
142. Devaux, J.J. (2014). [New insights on the organization of the nodes of Ranvier]. Revue neurologique 170, 819-824.
143. Devaux, J.J., Miura, Y., Fukami, Y., Inoue, T., Manso, C., Belghazi, M., Sekiguchi, K., Kokubun, N., Ichikawa, H., Wong, A.H., et al. (2016). Neurofascin-155 IgG4 in chronic inflammatory demyelinating polyneuropathy. Neurology 86, 800-807.
144. Kawamura, N., Yamasaki, R., Yonekawa, T., Matsushita, T., Kusunoki, S., Nagayama, S., Fukuda, Y., Ogata, H., Matsuse, D., Murai, H., et al. (2013). Anti-neurofascin antibody in patients with combined central and peripheral demyelination. Neurology 81, 714-722.
145. Cortese, A., Devaux, J.J., Zardini, E., Manso, C., Taieb, G., Carra Dalliere, C., Merle, P., Osera, C., Romagnolo, S., Visigalli, N., et al. (2016). Neurofascin-155 as a putative antigen in combined central and peripheral demyelination. Neurology(R) neuroimmunology & neuroinflammation 3, e238.
146. McKenzie, I.A., Ohayon, D., Li, H., de Faria, J.P., Emery, B., Tohyama, K., and Richardson, W.D. (2014). Motor skill learning requires active central myelination. Science (New York, NY) 346, 318-322.
147. Xiao, L., Ohayon, D., McKenzie, I.A., Sinclair-Wilson, A., Wright, J.L., Fudge, A.D., Emery, B., Li, H., and Richardson, W.D. (2016). Rapid production of new oligodendrocytes is required in the earliest stages of motor-skill learning. Nature neuroscience 19, 1210-1217.
148. Farwell Hagman, K.D., Shinde, D.N., Mroske, C., Smith, E., Radtke, K., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Alcaraz, W.A., et al. (2017). Candidate-gene criteria for clinical reporting: diagnostic exome sequencing identifies altered candidate genes among 8% of patients with undiagnosed diseases. Genet Med 19, 224-235.
149. Stoodley, C.J. (2016). The Cerebellum and Neurodevelopmental Disorders. Cerebellum 15, 34-37.
150. Nachtergaele, S., and He, C. (2017). The emerging biology of RNA post-transcriptional modifications. RNA Biol 14, 156-163.
151. Shav-Tal, Y., and Zipori, D. (2002). PSF and p54(nrb)/NonO--multi-functional nuclear proteins. FEBS letters 531, 109-114.
152. Bottini, S., Hamouda-Tekaya, N., Mategot, R., Zaragosi, L.E., Audebert, S., Pisano, S., Grandjean, V., Mauduit, C., Benahmed, M., Barbry, P., et al. (2017). Post-transcriptional gene silencing mediated by microRNAs is controlled by nucleoplasmic Sfpq. Nat Commun 8, 1189.
153. Takeuchi, A., Iida, K., Tsubota, T., Hosokawa, M., Denawa, M., Brown, J.B., Ninomiya, K., Ito, M., Kimura, H., Abe, T., et al. (2018). Loss of Sfpq Causes Long-Gene Transcriptopathy in the Brain. Cell reports 23, 1326-1341.
154. Cosker, K.E., Fenstermacher, S.J., Pazyra-Murphy, M.F., Elliott, H.L., and Segal, R.A. (2016). The RNA-binding protein SFPQ orchestrates an RNA regulon to promote axon viability. Nature neuroscience 19, 690-696.
155. Ke, Y.D., Dramiga, J., Schutz, U., Kril, J.J., Ittner, L.M., Schroder, H., and Gotz, J. (2012). Tau-mediated nuclear depletion and cytoplasmic accumulation of SFPQ in Alzheimer's and Pick's disease. PLoS One 7, e35678.
156. Stamova, B.S., Tian, Y., Nordahl, C.W., Shen, M.D., Rogers, S., Amaral, D.G., and Sharp, F.R. (2013). Evidence for differential alternative splicing in blood of young boys with autism spectrum disorders. Molecular autism 4, 30.
157. Tapia-Paez, I., Tammimies, K., Massinen, S., Roy, A.L., and Kere, J. (2008). The complex of TFII-I, PARP1, and SFPQ proteins regulates the DYX1C1 gene implicated in neuronal migration and dyslexia. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 22, 3001-3009.
158. Chang, J., Gilman, S.R., Chiang, A.H., Sanders, S.J., and Vitkup, D. (2015). Genotype to phenotype relationships in autism spectrum disorders. Nature neuroscience 18, 191-198.
159. Thomas-Jinu, S., Gordon, P.M., Fielding, T., Taylor, R., Smith, B.N., Snowden, V., Blanc, E., Vance, C., Topp, S., Wong, C.H., et al. (2017). Non-nuclear Pool of Splicing Factor SFPQ Regulates Axonal Transcripts Required for Normal Motor Development. Neuron 94, 931.
160. Passon, D.M., Lee, M., Rackham, O., Stanley, W.A., Sadowska, A., Filipovska, A., Fox, A.H., and Bond, C.S. (2012). Structure of the heterodimer of human NONO and paraspeckle protein component 1 and analysis of its role in subnuclear body formation. Proc Natl Acad Sci U S A 109, 4846-4850.

161. Luisier, R., Tyzack, G.E., Hall, C.E., Mitchell, J.S., Devine, H., Taha, D.M., Malik, B., Meyer, I., Greensmith, L., Newcombe, J., et al. (2018). Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. Nat Commun 9, 2010.
162. Takayama, K.I., Fujiwara, K., and Inoue, S. (2019). Amyloid precursor protein, an androgen-regulated gene, is targeted by RNA-binding protein PSF/SFPQ in neuronal cells. Genes Cells.
163. Ishigaki, S., Fujioka, Y., Okada, Y., Riku, Y., Udagawa, T., Honda, D., Yokoi, S., Endo, K., Ikenaka, K., Takagi, S., et al. (2017). Altered Tau Isoform Ratio Caused by Loss of FUS and SFPQ Function Leads to FTLD-like Phenotypes. Cell reports 18, 1118-1131.
164. Gordon, P.M., Efthymiou, S., Salpietro, V., Fielding, T., Borgione, E., Scuderi, C., Houlden, H., and Houart, C. (2020). Human patient SFPQ homozygous mutation is found deleterious for brain and motor development in a zebrafish model. bioRxiv.
165. Hotchkiss, L., Donkervoort, S., Leach, M.E., Mohassel, P., Bharucha-Goebel, D.X., Bradley, N., Nguyen, D., Hu, Y., Gurgel-Giannetti, J., and 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, 1114-1119.
166. Cherot, E., Keren, B., Dubourg, C., Carre, W., Fradin, M., Lavillaureix, A., Afenjar, A., Burglen, L., Whalen, S., Charles, P., et al. (2018). Using medical exome sequencing to identify the causes of neurodevelopmental disorders: Experience of 2 clinical units and 216 patients. Clin Genet 93, 567-576.
167. Ohba, C., Haginoya, K., Osaka, H., Kubota, K., Ishiyama, A., Hiraide, T., Komaki, H., Sasaki, M., Miyatake, S., Nakashima, M., et al. (2015). De novo KIF1A mutations cause intellectual deficit, cerebellar atrophy, lower limb spasticity and visual disturbance. J Hum Genet 60, 739-742.
168. Hosokawa, M., Takeuchi, A., Tanihata, J., Iida, K., Takeda, S., and Hagiwara, M. (2019). Loss of RNA-Binding Protein Sfpq Causes Long-Gene Transcriptopathy in Skeletal Muscle and Severe Muscle Mass Reduction with Metabolic Myopathy. iScience 13, 229-242.
169. Stiles, J. (2011). Brain development and the nature versus nurture debate. Prog Brain Res 189, 3-22.
170. Stiles, J., and Jernigan, T.L. (2010). The basics of brain development. Neuropsychol Rev 20, 327-348.
171. Sur, M., and Rubenstein, J.L. (2005). Patterning and plasticity of the cerebral cortex. Science (New York, NY) 310, 805-810.
172. Bishop, K.M., Rubenstein, J.L., and O'Leary, D.D. (2002). Distinct actions of Emx1, Emx2, and Pax6 in regulating the specification of areas in the developing neocortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 22, 7627-7638.
173. Tau, G.Z., and Peterson, B.S. (2010). Normal development of brain circuits. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology 35, 147-168.
174. Arora, N.K., Nair, M.K.C., Gulati, S., Deshmukh, V., Mohapatra, A., Mishra, D., Patel, V., Pandey, R.M., Das, B.C., Divan, G., et al. (2018). Neurodevelopmental disorders in children aged 2-9 years: Population-based burden estimates across five regions in India. PLoS Med 15, e1002615.
175. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. Nature 515, 216-221.
176. Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell 146, 247-261.
177. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209-215.
178. Ilyas, M., Mir, A., Efthymiou, S., and Houlden, H. (2020). The genetics of intellectual disability: advancing technology and gene editing. F1000Res 9.
179. Perera, F., and Herbstman, J. (2011). Prenatal environmental exposures, epigenetics, and disease. Reprod Toxicol 31, 363-373.
180. Hoischen, A., van Bon, B.W., Rodriguez-Santiago, B., Gilissen, C., Vissers, L.E., de Vries, P., Janssen, I., van Lier, B., Hastings, R., Smithson, S.F., et al. (2011). De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. Nature genetics 43, 729-731.
181. Zhang, P., Xing, C., Rhodes, S.D., He, Y., Deng, K., Li, Z., He, F., Zhu, C., Nguyen, L., Zhou, Y., et al. (2016). Loss of Asxl1 Alters Self-Renewal and Cell Fate of Bone Marrow Stromal Cell, Leading to Bohring-Opitz-like Syndrome in Mice. Stem cell reports 6, 914-925.
182. Arunachal, G., Danda, S., Omprakash, S., and Kumar, S. (2016). A novel de-novo frameshift mutation of the ASXL1 gene in a classic case of Bohring-Opitz syndrome. Clinical dysmorphology 25, 101-105.
183. Bedoukian, E., Copenheaver, D., Bale, S., and Deardorff, M. (2018). Bohring-Opitz syndrome caused by an ASXL1 mutation inherited from a germline mosaic mother. American journal of medical genetics Part A 176, 1249-1252.
184. Carlston, C.M., O'Donnell-Luria, A.H., Underhill, H.R., Cummings, B.B., Weisburd, B., Minikel, E.V., Birnbaum, D.P., Tvrdik, T., MacArthur, D.G., and Mao, R. (2017). Pathogenic ASXL1 somatic variants in reference databases complicate germline variant interpretation for Bohring-Opitz Syndrome. Human mutation 38, 517-523.
185. Dangiolo, S.B., Wilson, A., Jobanputra, V., and Anyane-Yeboa, K. (2015). Bohring-Opitz syndrome (BOS) with a new ASXL1 pathogenic variant: Review of the most prevalent molecular and phenotypic features of the syndrome. American journal of medical genetics Part A 167a, 3161-3166.
186. Magini, P., Della Monica, M., Uzielli, M.L., Mongelli, P., Scarselli, G., Gambineri, E., Scarano, G., and Seri, M. (2012). Two novel patients with Bohring-Opitz syndrome caused by de novo ASXL1 mutations. American journal of medical genetics Part A 158a, 917-921.
187. Urreizti, R., Roca-Ayats, N., Trepat, J., Garcia-Garcia, F., Aleman, A., Orteschi, D., Marangi, G., Neri, G., Opitz, J.M., Dopazo, J., et al. (2016). Screening of CD96 and ASXL1 in 11 patients with Opitz C or Bohring-Opitz syndromes. American journal of medical genetics Part A 170a, 24-31.

188. Efthymiou, S., Salpietro, V., Pironti, E., Bonsignore, M., Ferrazzoli, V., Rosa, G.D., and Houlden, H. (2019). A de novo truncating mutation in ASXL1 associated with segmental overgrowth. J Genet 98.
189. Russell, B., Johnston, J.J., Biesecker, L.G., Kramer, N., Pickart, A., Rhead, W., Tan, W.H., Brownstein, C.A., Kate Clarkson, L., Dobson, A., et al. (2015). Clinical management of patients with ASXL1 mutations and Bohring-Opitz syndrome, emphasizing the need for Wilms tumor surveillance. American journal of medical genetics Part A 167a, 2122-2131.
190. Malik, S. (2014). Polydactyly: phenotypes, genetics and classification. Clin Genet 85, 203-212.
191. Bruel, A.L., Franco, B., Duffourd, Y., Thevenon, J., Jego, L., Lopez, E., Deleuze, J.F., Doummar, D., Giles, R.H., Johnson, C.A., et al. (2017). Fifteen years of research on oral-facial-digital syndromes: from 1 to 16 causal genes. Journal of medical genetics 54, 371-380.
192. Castilla, E.E., Lugarinho, R., da Graca Dutra, M., and Salgado, L.J. (1998). Associated anomalies in individuals with polydactyly. Am J Med Genet 80, 459-465.
193. Oliver, C.P. (1940). Recessive polydactylism associated with mental deficiency. J Hered 31, 365:367.
194. R. E. Stevenson, G.W. (1983). Polydactyly and mental retardation in siblings. Proc Greenwood Genet Center 2, 20-22.
195. Salpietro, C.D., Briuglia, S., Bertuccio, G., Rigoli, L., Mingarelli, R., and Dallapiccola, B. (2005). Report of a third family with Oliver syndrome. American journal of medical genetics Part A 139A, 159-161.
196. Salpietro, V., Efthymiou, S., Manole, A., Maurya, B., Wiethoff, S., Ashokkumar, B., Cutrupi, M.C., Dipasquale, V., Manti, S., Botia, J.A., et al. (2018). A loss-of-function homozygous mutation in DDX59 implicates a conserved DEAD-box RNA helicase in nervous system development and function. Human mutation 39, 187-192.
197. Shamseldin, H.E., Rajab, A., Alhashem, A., Shaheen, R., Al-Shidi, T., Alamro, R., Al Harassi, S., and Alkuraya, F.S. (2013). Mutations in DDX59 implicate RNA helicase in the pathogenesis of orofaciodigital syndrome. Am J Hum Genet 93, 555-560.
198. Rocak, S., and Linder, P. (2004). DEAD-box proteins: the driving forces behind RNA metabolism. Nat Rev Mol Cell Biol 5, 232-241.
199. Jarmoskaite, I., and Russell, R. (2014). RNA helicase proteins as chaperones and remodelers. Annu Rev Biochem 83, 697-725.
200. Faily, S., Perveen, R., Urquhart, J., Chandler, K., and Clayton-Smith, J. (2017). Confirmation that mutations in DDX59 cause an autosomal recessive form of oral-facial-digital syndrome: Further delineation of the DDX59 phenotype in two new families. European journal of medical genetics 60, 527-532.
201. Metin, C., and Pedraza, M. (2014). Cilia: traffic directors along the road of cortical development. The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry 20, 468-482.
202. Bettencourt, C., Forabosco, P., Wiethoff, S., Heidari, M., Johnstone, D.M., Botia, J.A., Collingwood, J.F., Hardy, J., Consortium, U.K.B.E., Milward, E.A., et al. (2016). Gene co-expression networks shed light into diseases of brain iron accumulation. Neurobiol Dis 87, 59-68.
203. Lee, J.E., and Gleeson, J.G. (2011). Cilia in the nervous system: linking cilia function and neurodevelopmental disorders. Current opinion in neurology 24, 98-105.
204. Tan, P.L., Barr, T., Inglis, P.N., Mitsuma, N., Huang, S.M., Garcia-Gonzalez, M.A., Bradley, B.A., Coforio, S., Albrecht, P.J., Watnick, T., et al. (2007). Loss of Bardet Biedl syndrome proteins causes defects in peripheral sensory innervation and function. Proc Natl Acad Sci U S A 104, 17524-17529.
205. Jang, M.A., Kim, E.K., Now, H., Nguyen, N.T., Kim, W.J., Yoo, J.Y., Lee, J., Jeong, Y.M., Kim, C.H., Kim, O.H., et al. (2015). Mutations in DDX58, which encodes RIG-I, cause atypical Singleton-Merten syndrome. Am J Hum Genet 96, 266-274.
206. Di Donato, I., Bianchi, S., De Stefano, N., Dichgans, M., Dotti, M.T., Duering, M., Jouvent, E., Korczyn, A.D., Lesnik-Oberstein, S.A., Malandrini, A., et al. (2017). Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL) as a model of small vessel disease: update on clinical, diagnostic, and management aspects. BMC Med 15, 41.
207. Surabhi, S., Tripathi, B.K., Maurya, B., Bhaskar, P.K., Mukherjee, A., and Mutsuddi, M. (2015). Regulation of Notch Signaling by an Evolutionary Conserved DEAD Box RNA Helicase, Maheshvara in Drosophila melanogaster. Genetics 201, 1071-1085.
208. Lee, E.Y., Kim, S., and Kim, M.H. (2018). Aminoacyl-tRNA synthetases, therapeutic targets for infectious diseases. Biochem Pharmacol 154, 424-434.
209. Ognjenovic, J., and Simonovic, M. (2018). Human aminoacyl-tRNA synthetases in diseases of the nervous system. RNA Biol 15, 623-634.
210. Rajendran, V., Kalita, P., Shukla, H., Kumar, A., and Tripathi, T. (2018). Aminoacyl-tRNA synthetases: Structure, function, and drug discovery. Int J Biol Macromol 111, 400-414.
211. Antonellis, A., and Green, E.D. (2008). The role of aminoacyl-tRNA synthetases in genetic diseases. Annu Rev Genomics Hum Genet 9, 87-107.
212. Meyer-Schuman, R., and Antonellis, A. (2017). Emerging mechanisms of aminoacyl-tRNA synthetase mutations in recessive and dominant human disease. Hum Mol Genet 26, R114-R127.
213. Francklyn, C.S., and Mullen, P. (2019). Progress and Challenges in Aminoacyl-tRNA Synthetase-based Therapeutics. J Biol Chem.
214. Rogers, S.O. (2019). Evolution of the genetic code based on conservative changes of codons, amino acids, and aminoacyl tRNA synthetases. J Theor Biol 466, 1-10.
215. Gonzalez-Serrano, L.E., Chihade, J.W., and Sissler, M. (2019). When a common biological role does not imply common disease outcomes: Disparate pathology linked to human mitochondrial aminoacyl-tRNA synthetases. J Biol Chem.
216. Okur, V., Ganapathi, M., Wilson, A., and Chung, W.K. (2018). Biallelic variants in VARS in a family with two siblings with intellectual disability and microcephaly: case report and review of the literature. Cold Spring Harb Mol Case Stud 4.
217. Stephen, J., Nampoothiri, S., Banerjee, A., Tolman, N.J., Penninger, J.M., Elling, U., Agu, C.A., Burke, J.D., Devadathan, K., Kannan, R., et al. (2018). Loss of function mutations in VARS encoding cytoplasmic valyl-tRNA synthetase cause microcephaly, seizures, and progressive cerebral atrophy. Hum Genet 137, 293-303.

218. Siekierska, A., Stamberger, H., Deconinck, T., Oprescu, S.N., Partoens, M., Zhang, Y., Sourbron, J., Adriaenssens, E., Mullen, P., Wiencek, P., et al. (2019). Biallelic VARS variants cause developmental encephalopathy with microcephaly that is recapitulated in vars knockout zebrafish. Nat Commun 10, 708.
219. Friedman, J., Smith, D.E., Issa, M.Y., Stanley, V., Wang, R., Mendes, M.I., Wright, M.S., Wigby, K., Hildreth, A., Crawford, J.R., et al. (2019). Biallelic mutations in valyl-tRNA synthetase gene VARS are associated with a progressive neurodevelopmental epileptic encephalopathy. Nat Commun 10, 707.
220. Krenke, K., Szczaluba, K., Bielecka, T., Rydzanicz, M., Lange, J., Koppolu, A., and Ploski, R. (2019). FARSA mutations mimic phenylalanyl-tRNA synthetase deficiency caused by FARSB defects. Clin Genet 96, 468-472.
221. Forrester, N., Rattihalli, R., Horvath, R., Maggi, L., Manzur, A., Fuller, G., Gutowski, N., Rankin, J., Dick, D., Buxton, C., et al. (2020). Clinical and Genetic Features in a Series of Eight Unrelated Patients with Neuropathy Due to Glycyl-tRNA Synthetase (GARS) Variants. J Neuromuscul Dis 7, 137-143.
222. Lee, A.J., Nam, D.E., Choi, Y.J., Nam, S.H., Choi, B.O., and Chung, K.W. (2020). Alanyl-tRNA synthetase 1 (AARS1) gene mutation in a family with intermediate Charcot-Marie-Tooth neuropathy. Genes Genomics.
223. Williams, K.B., Brigatti, K.W., Puffenberger, E.G., Gonzaga-Jauregui, C., Griffin, L.B., Martinez, E.D., Wenger, O.K., Yoder, M.A., Kandula, V.V.R., Fox, M.D., et al. (2019). Homozygosity for a mutation affecting the catalytic domain of tyrosyl-tRNA synthetase (YARS) causes multisystem disease. Hum Mol Genet 28, 525-538.
224. Park, J.S., Park, M.C., Lee, K.Y., Goughnour, P.C., Jeong, S.J., Kim, H.S., Kim, H.J., Lee, B.J., Kim, S., and Han, B.W. (2018). Unique N-terminal extension domain of human asparaginyl-tRNA synthetase elicits CCR3-mediated chemokine activity. Int J Biol Macromol 120, 835-845.
225. Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., and Hinnebusch, A.G. (2000). Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. Mol Cell 6, 269-279.
226. Vijayakumar, R., and Tripathi, T. (2018). Soluble expression and purification of a full-length asparaginyl tRNA synthetase from Fasciola gigantica. Protein Expr Purif 143, 9-13.
227. He, W., Bai, G., Zhou, H., Wei, N., White, N.M., Lauer, J., Liu, H., Shi, Y., Dumitru, C.D., Lettieri, K., et al. (2015). CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. Nature 526, 710-714.
228. Boczonadi, V., Meyer, K., Gonczarowska-Jorge, H., Griffin, H., Roos, A., Bartsakoulia, M., Bansagi, B., Ricci, G., Palinkas, F., Zahedi, R.P., et al. (2018). Mutations in glycyl-tRNA synthetase impair mitochondrial metabolism in neurons. Hum Mol Genet 27, 2187-2204.
229. McClain, W.H., Schneider, J., Bhattacharya, S., and Gabriel, K. (1998). The importance of tRNA backbone-mediated interactions with synthetase for aminoacylation. Proc Natl Acad Sci U S A 95, 460-465.
230. Bicknell, L.S., Bongers, E.M., Leitch, A., Brown, S., Schoots, J., Harley, M.E., Aftimos, S., Al-Aama, J.Y., Bober, M., Brown, P.A., et al. (2011). Mutations in the pre-replication complex cause Meier-Gorlin syndrome. Nature genetics 43, 356-359.
231. Schmelzer, L., Smitka, M., Wolf, C., Lucas, N., Tungler, V., Hahn, G., Tzschach, A., Di Donato, N., Lee-Kirsch, M.A., and von der Hagen, M. (2018). Variable clinical phenotype in two siblings with Aicardi-Goutieres syndrome type 6 and a novel mutation in the ADAR gene. European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society 22, 186-189.
232. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10, 845-858.
233. Mele, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al. (2015). Human genomics. The human transcriptome across tissues and individuals. Science 348, 660-665.
234. Consortium, G.T. (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348, 648-660.
235. Seaver, L.H., DeRoos, S., Betz, B., and Rajasekaran, S. (2019). Reply to Finsterer Regarding Lethal NARS2-Related Disorder Associated With Rapidly Progressive Intractable Epilepsy and Global Brain Atrophy. Pediatric neurology 93, 65.
236. Simon, M., Richard, E.M., Wang, X., Shahzad, M., Huang, V.H., Qaiser, T.A., Potluri, P., Mahl, S.E., Davila, A., Nazli, S., et al. (2015). Mutations of human NARS2, encoding the mitochondrial asparaginyl-tRNA synthetase, cause nonsyndromic deafness and Leigh syndrome. PLoS Genet 11, e1005097.
237. Sofou, K., Kollberg, G., Holmstrom, M., Davila, M., Darin, N., Gustafsson, C.M., Holme, E., Oldfors, A., Tulinius, M., and Asin-Cayuela, J. (2015). Whole exome sequencing reveals mutations in NARS2 and PARS2, encoding the mitochondrial asparaginyl-tRNA synthetase and prolyl-tRNA synthetase, in patients with Alpers syndrome. Molecular genetics & genomic medicine 3, 59-68.
238. Vanlander, A.V., Menten, B., Smet, J., De Meirleir, L., Sante, T., De Paepe, B., Seneca, S., Pearce, S.F., Powell, C.A., Vergult, S., et al. (2015). Two siblings with homozygous pathogenic splice-site variant in mitochondrial asparaginyl-tRNA synthetase (NARS2). Human mutation 36, 222-231.
239. Boczonadi, V., Jennings, M.J., and Horvath, R. (2018). The role of tRNA synthetases in neurological and neuromuscular disorders. FEBS Lett 592, 703-717.
240. Taylor, W.R. (1986). The classification of amino acid conservation. J Theor Biol 119, 205-218.
241. Taylor, W.R. (1986). The classification of amino acid conservation. Journal of theoretical Biology 119, 205-218.
242. Theil, A.F., Botta, E., Raams, A., Smith, D.E.C., Mendes, M.I., Caligiuri, G., Giachetti, S., Bione, S., Carriero, R., Liberi, G., et al. (2019). Bi-allelic TARS Mutations Are Associated with Brittle Hair Phenotype. Am J Hum Genet 105, 434-440.
243. Jeong, S.J., Park, S., Nguyen, L.T., Hwang, J., Lee, E.Y., Giong, H.K., Lee, J.S., Yoon, I., Lee, J.H., Kim, J.H., et al. (2019). A threonyl-tRNA synthetase-mediated translation initiation machinery. Nat Commun 10, 1357.
244. Wellman, T.L., Eckenstein, M., Wong, C., Rincon, M., Ashikaga, T., Mount, S.L., Francklyn, C.S., and Lounsbury, K.M. (2014). Threonyl-tRNA synthetase overexpression correlates with angiogenic markers and progression of human ovarian cancer. BMC Cancer 14, 620.

245. Williams, T.F., Mirando, A.C., Wilkinson, B., Francklyn, C.S., and Lounsbury, K.M. (2013). Secreted Threonyl-tRNA synthetase stimulates endothelial cell migration and angiogenesis. Scientific reports 3, 1317.
246. Mirando, A.C., Fang, P., Williams, T.F., Baldor, L.C., Howe, A.K., Ebert, A.M., Wilkinson, B., Lounsbury, K.M., Guo, M., and Francklyn, C.S. (2015). Aminoacyl-tRNA synthetase dependent angiogenesis revealed by a bioengineered macrolide inhibitor. Scientific reports 5, 13160.
247. Guo, M., and Schimmel, P. (2013). Essential nontranslational functions of tRNA synthetases. Nat Chem Biol 9, 145-153.
248. Myers, K.A., Johnstone, D.L., and Dyment, D.A. (2019). Epilepsy genetics: Current knowledge, applications, and future directions. Clin Genet 95, 95-111.
249. Salpietro, V., Dixon, C.L., Guo, H., Bello, O.D., Vandrovcova, J., Efthymiou, S., Maroofian, R., Heimer, G., Burglen, L., Valence, S., et al. (2019). AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. Nat Commun 10, 3094.
250. Tham, E., Lindstrand, A., Santani, A., Malmgren, H., Nesbitt, A., Dubbs, H.A., Zackai, E.H., Parker, M.J., Millan, F., Rosenbaum, K., et al. (2015). Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. Am J Hum Genet 96, 507-513.
251. Murray, C.R., Abel, S.N., McClure, M.B., Foster, J., 2nd, Walke, M.I., Jayakar, P., Bademci, G., and Tekin, M. (2017). Novel Causative Variants in DYRK1A, KARS, and KAT6A Associated with Intellectual Disability and Additional Phenotypic Features. Journal of pediatric genetics 6, 77-83.
252. Millan, F., Cho, M.T., Retterer, K., Monaghan, K.G., Bai, R., Vitazka, P., Everman, D.B., Smith, B., Angle, B., Roberts, V., et al. (2016). Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. American journal of medical genetics Part A 170, 1791-1798.
253. Arboleda, V.A., Lee, H., Dorrani, N., Zadeh, N., Willis, M., Macmurdo, C.F., Manning, M.A., Kwan, A., Hudgins, L., Barthelemy, F., et al. (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, 498-506.
254. Zwaveling-Soonawala, N., Alders, M., Jongejan, A., Kovacic, L., Duijkers, F.A., Maas, S.M., Fliers, E., van Trotsenburg, A.S.P., and Hennekam, R.C. (2018). Clues for Polygenic Inheritance of Pituitary Stalk Interruption Syndrome From Exome Sequencing in 20 Patients. The Journal of clinical endocrinology and metabolism 103, 415-428.
255. Good-Jacobson, K.L., Chen, Y., Voss, A.K., Smyth, G.K., Thomas, T., and Tarlinton, D. (2014). Regulation of germinal center responses and B-cell memory by the chromatin modifier MOZ. Proc Natl Acad Sci U S A 111, 9585-9590.
256. Voss, A.K., Vanyai, H.K., Collin, C., Dixon, M.P., McLennan, T.J., Sheikh, B.N., Scambler, P., and Thomas, T. (2012). MOZ regulates the Tbx1 locus, and Moz mutation partially phenocopies DiGeorge syndrome. Dev Cell 23, 652-663.
257. Yan, K., Rousseau, J., Littlejohn, R.O., Kiss, C., Lehman, A., Rosenfeld, J.A., Stumpel, C.T., Stegmann, A.P., Robak, L., Scaglia, F., et al. (2017). Mutations in the Chromatin Regulator Gene BRPF1 Cause Syndromic Intellectual Disability and Deficient Histone Acetylation. Am J Hum Genet 100, 91-104.
258. Dreissen, Y.E.M., Boeree, T., Koelman, J., and Tijssen, M.A.J. (2017). Startle responses in functional jerky movement disorders are increased but have a normal pattern. Parkinsonism & related disorders 40, 27-32.
259. Pagnamenta, A.T., Murakami, Y., Taylor, J.M., Anzilotti, C., Howard, M.F., Miller, V., Johnson, D.S., Tadros, S., Mansour, S., Temple, I.K., et al. (2017). Analysis of exome data for 4293 trios suggests GPI-anchor biogenesis defects are a rare cause of developmental disorders. European journal of human genetics : EJHG 25, 669-679.
260. Kinoshita, T., and Fujita, M. (2016). Biosynthesis of GPI-anchored proteins: special emphasis on GPI lipid remodeling. J Lipid Res 57, 6-24.
261. McKean, D.M., and Niswander, L. (2012). Defects in GPI biosynthesis perturb Cripto signaling during forebrain development in two new mouse models of holoprosencephaly. Biol Open 1, 874-883.
262. Park, S., Lee, C., Sabharwal, P., Zhang, M., Meyers, C.L., and Sockanathan, S. (2013). GDE2 promotes neurogenesis by glycosylphosphatidylinositol-anchor cleavage of RECK. Science (New York, NY) 339, 324-328.
263. Allegri, G., Bertazzo, A., Biasiolo, M., Costa, C.V., and Ragazzi, E. (2003). Kynurenine pathway enzymes in different species of animals. Adv Exp Med Biol 527, 455-463.
264. Nozaki, M., Ohishi, K., Yamada, N., Kinoshita, T., Nagy, A., and Takeda, J. (1999). Developmental abnormalities of glycosylphosphatidylinositol-anchor-deficient embryos revealed by Cre/loxP system. Lab Invest 79, 293-299.
265. Alfieri, J.A., Martin, A.D., Takeda, J., Kondoh, G., Myles, D.G., and Primakoff, P. (2003). Infertility in female mice with an oocyte-specific knockout of GPI-anchored proteins. Journal of cell science 116, 2149-2155.
266. Kondoh, G., Tojo, H., Nakatani, Y., Komazawa, N., Murata, C., Yamagata, K., Maeda, Y., Kinoshita, T., Okabe, M., Taguchi, R., et al. (2005). Angiotensin-converting enzyme is a GPI-anchored protein releasing factor crucial for fertilization. Nat Med 11, 160-166.
267. Um, J.W., and Ko, J. (2017). Neural Glycosylphosphatidylinositol-Anchored Proteins in Synaptic Specification. Trends Cell Biol 27, 931-945.
268. Scheffer, I.E., and Liao, J. (2020). Deciphering the concepts behind "Epileptic encephalopathy" and "Developmental and epileptic encephalopathy". European journal of paediatric neurology : EJPN : official journal of the European Paediatric Neurology Society 24, 11-14.
269. Murakami, Y., Nguyen, T.T.M., Baratang, N., Raju, P.K., Knaus, A., Ellard, S., Jones, G., Lace, B., Rousseau, J., Ajeawung, N.F., et al. (2019). Mutations in PIGB Cause an Inherited GPI Biosynthesis Defect with an Axonal Neuropathy and Metabolic Abnormality in Severe Cases. Am J Hum Genet 105, 384-394.
270. Martin, H.C., Kim, G.E., Pagnamenta, A.T., Murakami, Y., Carvill, G.L., Meyer, E., Copley, R.R., Rimmer, A., Barcia, G., Fleming, M.R., et al. (2014). Clinical whole-genome sequencing in severe early-onset epilepsy reveals new genes and improves molecular diagnosis. Hum Mol Genet 23, 3200-3211.

271. Johnstone, D.L., Nguyen, T.T., Murakami, Y., Kernohan, K.D., Tetreault, M., Goldsmith, C., Doja, A., Wagner, J.D., Huang, L., Hartley, T., et al. (2017). Compound heterozygous mutations in the gene PIGP are associated with early infantile epileptic encephalopathy. Hum Mol Genet 26, 1706-1715.
272. Johnston, J.J., Gropman, A.L., Sapp, J.C., Teer, J.K., Martin, J.M., Liu, C.F., Yuan, X., Ye, Z., Cheng, L., Brodsky, R.A., et al. (2012). The phenotype of a germline mutation in PIGA: the gene somatically mutated in paroxysmal nocturnal hemoglobinuria. Am J Hum Genet 90, 295-300.
273. Ohishi, K., Inoue, N., and Kinoshita, T. (2001). PIG-S and PIG-T, essential for GPI anchor attachment to proteins, form a complex with GAA1 and GPI8. EMBO J 20, 4088-4098.
274. Nguyen, T.T.M., Murakami, Y., Wigby, K.M., Baratang, N.V., Rousseau, J., St-Denis, A., Rosenfeld, J.A., Laniewski, S.C., Jones, J., Iglesias, A.D., et al. (2018). Mutations in PIGS, Encoding a GPI Transamidase, Cause a Neurological Syndrome Ranging from Fetal Akinesia to Epileptic Encephalopathy. Am J Hum Genet 103, 602-611.
275. Zhang, L., Mao, X., Long, H., Xiao, B., Luo, Z., Xiao, W., and Jin, X. (2020). Compound Heterozygous PIGS Variants Associated With Infantile Spasm, Global Developmental Delay, Hearing Loss, Visual Impairment, and Hypotonia. Front Genet 11, 564.
276. Kuki, I., Takahashi, Y., Okazaki, S., Kawawaki, H., Ehara, E., Inoue, N., Kinoshita, T., and Murakami, Y. (2013). Vitamin B6-responsive epilepsy due to inherited GPI deficiency. Neurology 81, 1467-1469.
277. Joshi, C., Kolbe, D.L., Mansilla, M.A., Mason, S., Smith, R.J., and Campbell, C.A. (2016). Ketogenic diet - A novel treatment for early epileptic encephalopathy due to PIGA deficiency. Brain & development 38, 848-851.
278. Ko, J., Choii, G., and Um, J.W. (2015). The balancing act of GABAergic synapse organizers. Trends Mol Med 21, 256-268.
279. Ramamoorthi, K., and Lin, Y. (2011). The contribution of GABAergic dysfunction to neurodevelopmental disorders. Trends Mol Med 17, 452-462.
280. Fritschy, J.M., and Tyagarajan, S.K. (2017). GABAergic Synaptogenesis: A Case for Cooperation. Neuron 96, 709-711.
281. Aruga, J., Yokota, N., and Mikoshiba, K. (2003). Human SLITRK family genes: genomic organization and expression profiling in normal brain and brain tumor tissue. Gene 315, 87-94.
282. Li, J., Han, W., Pelkey, K.A., Duan, J., Mao, X., Wang, Y.X., Craig, M.T., Dong, L., Petralia, R.S., McBain, C.J., et al. (2017). Molecular Dissection of Neuroligin 2 and Slitrk3 Reveals an Essential Framework for GABAergic Synapse Development. Neuron 96, 808-826 e808.
283. Kabankin, A.S., Sinauridze, E.I., Lipets, E.N., and Ataullakhanov, F.I. (2019). Computer Design of Low-Molecular-Weight Inhibitors of Coagulation Factors. Biochemistry (Mosc) 84, 119-136.
284. de Groot, C., Floriou-Servou, A., Tsai, Y.C., Fruh, S., Kohler, M., Parkin, G., Schwerdel, C., Bosshard, G., Kaila, K., Fritschy, J.M., et al. (2017). RhoGEF9 splice isoforms influence neuronal maturation and synapse formation downstream of alpha2 GABAA receptors. PLoS Genet 13, e1007073.
285. Honda, K., Katzke, V.A., Husing, A., Okaya, S., Shoji, H., Onidani, K., Olsen, A., Tjonneland, A., Overvad, K., Weiderpass, E., et al. (2019). CA19-9 and apolipoprotein-A2 isoforms as detection markers for pancreatic cancer: a prospective evaluation. Int J Cancer 144, 1877-1887.
286. Kang, H., Han, K.A., Won, S.Y., Kim, H.M., Lee, Y.H., Ko, J., and Um, J.W. (2016). Slitrk Missense Mutations Associated with Neuropsychiatric Disorders Distinctively Impair Slitrk Trafficking and Synapse Formation. Front Mol Neurosci 9, 104.
287. Wang, L., Li, Z., Sievert, D., Smith, D.E.C., Mendes, M.I., Chen, D.Y., Stanley, V., Ghosh, S., Wang, Y., Kara, M., et al. (2020). Loss of NARS1 impairs progenitor proliferation in cortical brain organoids and leads to microcephaly. Nat Commun 11, 4038.
288. Harper, J.L., Wilson, T.E., and Mitchell, R.M. (2020). Case report of two children with auditory neuropathy spectrum disorder related to a neurofascin (NFASC) gene variant. Int J Pediatr Otorhinolaryngol 131, 109863.
289. Kvarnung, M., Shahsavani, M., Taylan, F., Moslem, M., Breeuwsma, N., Laan, L., Schuster, J., Jin, Z., Nilsson, D., Lieden, A., et al. (2019). Ataxia in Patients With Bi-Allelic NFASC Mutations and Absence of Full-Length NF186. Front Genet 10, 896.
290. Querol, L., Rojas-Garcia, R., Diaz-Manera, J., Barcena, J., Pardo, J., Ortega-Moreno, A., Sedano, M.J., Sero-Ballesteros, L., Carvajal, A., Ortiz, N., et al. (2015). Rituximab in treatment-resistant CIDP with antibodies against paranodal proteins. Neurology(R) neuroimmunology & neuroinflammation 2, e149.
291. Klingseisen, A., Ristoiu, A.M., Kegel, L., Sherman, D.L., Rubio-Brotons, M., Almeida, R.G., Koudelka, S., Benito-Kwiecinski, S.K., Poole, R.J., Brophy, P.J., et al. (2019). Oligodendrocyte Neurofascin Independently Regulates Both Myelin Targeting and Sheath Growth in the CNS. Dev Cell 51, 730-744 e736.
292. Hinman, J.D., Peters, A., Cabral, H., Rosene, D.L., Hollander, W., Rasband, M.N., and Abraham, C.R. (2006). Age-related molecular reorganization at the node of Ranvier. The Journal of comparative neurology 495, 351-362.
293. Howell, O.W., Palser, A., Polito, A., Melrose, S., Zonta, B., Scheiermann, C., Vora, A.J., Brophy, P.J., and Reynolds, R. (2006). Disruption of neurofascin localization reveals early changes preceding demyelination and remyelination in multiple sclerosis. Brain 129, 3173-3185.
294. Low, K.J., Stals, K., Caswell, R., Wakeling, M., Clayton-Smith, J., Donaldson, A., Foulds, N., Norman, A., Splitt, M., Urankar, K., et al. (2018). Phenotype of CNTNAP1: a study of patients demonstrating a specific severe congenital hypomyelinating neuropathy with survival beyond infancy. European journal of human genetics : EJHG 26, 796-807.
295. Kagiava, A., Karaiskos, C., Richter, J., Tryfonos, C., Lapathitis, G., Sargiannidou, I., Christodoulou, C., and Kleopa, K.A. (2018). Intrathecal gene therapy in mouse models expressing CMT1X mutations. Hum Mol Genet 27, 1460-1473.
296. Kagiava, A., Richter, J., Tryfonos, C., Karaiskos, C., Heslegrave, A.J., Sargiannidou, I., Rossor, A.M., Zetterberg, H., Reilly, M.M., Christodoulou, C., et al. (2019). Gene replacement therapy after neuropathy onset provides therapeutic benefit in a model of CMT1X. Hum Mol Genet 28, 3528-3542.
297. Sargiannidou, I., Kagiava, A., and Kleopa, K.A. (2020). Gene therapy approaches targeting Schwann cells for demyelinating neuropathies. Brain Res 1728, 146572.

298. Latour, P., Thauvin-Robinet, C., Baudelet-Mery, C., Soichot, P., Cusin, V., Faivre, L., Locatelli, M.C., Mayencon, M., Sarcey, A., Broussolle, E., et al. (2010). A major determinant for binding and aminoacylation of tRNA(Ala) in cytoplasmic Alanyl-tRNA synthetase is mutated in dominant axonal Charcot-Marie-Tooth disease. Am J Hum Genet 86, 77-82.
299. Seburn, K.L., Nangle, L.A., Cox, G.A., Schimmel, P., and Burgess, R.W. (2006). An active dominant mutation of glycyl-tRNA synthetase causes neuropathy in a Charcot-Marie-Tooth 2D mouse model. Neuron 51, 715-726.
300. Lynch, D.S., Zhang, W.J., Lakshmanan, R., Kinsella, J.A., Uzun, G.A., Karbay, M., Tufekcioglu, Z., Hanagasi, H., Burke, G., Foulds, N., et al. (2016). Analysis of Mutations in AARS2 in a Series of CSF1R-Negative Patients With Adult-Onset Leukoencephalopathy With Axonal Spheroids and Pigmented Glia. JAMA Neurol 73, 1433-1439.
301. Dohrn, M.F., Glockle, N., Mulahasanovic, L., Heller, C., Mohr, J., Bauer, C., Riesch, E., Becker, A., Battke, F., Hortnagel, K., et al. (2017). Frequent genes in rare diseases: panel-based next generation sequencing to disclose causal mutations in hereditary neuropathies. J Neurochem 143, 507-522.
302. Sundal, C., Carmona, S., Yhr, M., Almstrom, O., Ljungberg, M., Hardy, J., Hedberg-Oldfors, C., Fred, A., Bras, J., Oldfors, A., et al. (2019). An AARS variant as the likely cause of Swedish type hereditary diffuse leukoencephalopathy with spheroids. Acta Neuropathol Commun 7, 188.
303. Dallabona, C., Diodato, D., Kevelam, S.H., Haack, T.B., Wong, L.J., Salomons, G.S., Baruffini, E., Melchionda, L., Mariotti, C., Strom, T.M., et al. (2014). Novel (ovario) leukodystrophy related to AARS2 mutations. Neurology 82, 2063-2071.
304. Gotz, A., Tyynismaa, H., Euro, L., Ellonen, P., Hyotylainen, T., Ojala, T., Hamalainen, R.H., Tommiska, J., Raivio, T., Oresic, M., et al. (2011). Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. Am J Hum Genet 88, 635-642.
305. Sommerville, E.W., Zhou, X.L., Olahova, M., Jenkins, J., Euro, L., Konovalova, S., Hilander, T., Pyle, A., He, L., Habeebu, S., et al. (2019). Instability of the mitochondrial alanyl-tRNA synthetase underlies fatal infantile-onset cardiomyopathy. Hum Mol Genet 28, 258-268.
306. Wang, X., Wang, Q., Tang, H., Chen, B., Dong, X., Niu, S., Li, S., Shi, Y., Shan, W., and Zhang, Z. (2019). Novel Alanyl-tRNA Synthetase 2 Pathogenic Variants in Leukodystrophies. Frontiers in neurology 10, 1321.
307. Kuo, M.E., Antonellis, A., and Shakkottai, V.G. (2020). Alanyl-tRNA Synthetase 2 (AARS2)-Related Ataxia Without Leukoencephalopathy. Cerebellum 19, 154-160.
308. Kuo, M.E., Theil, A.F., Kievit, A., Malicdan, M.C., Introne, W.J., Christian, T., Verheijen, F.W., Smith, D.E.C., Mendes, M.I., Hussaarts-Odijk, L., et al. (2019). Cysteinyl-tRNA Synthetase Mutations Cause a Multi-System, Recessive Disease That Includes Microcephaly, Developmental Delay, and Brittle Hair and Nails. Am J Hum Genet 104, 520-529.
309. Coughlin, C.R., 2nd, Scharer, G.H., Friederich, M.W., Yu, H.C., Geiger, E.A., Creadon-Swindell, G., Collins, A.E., Vanlander, A.V., Coster, R.V., Powell, C.A., et al. (2015). Mutations in the mitochondrial cysteinyl-tRNA synthase gene, CARS2, lead to a severe epileptic encephalopathy and complex movement disorder. J Med Genet 52, 532-540.
310. Samanta, D., Gokden, M., and Willis, E. (2018). Clinicopathologic Findings of CARS2 Mutation. Pediatr Neurol 87, 65-69.
311. Samanta, D. (2018). Cerebral Infarction in CARS2 Mutation. Pediatr Neurol.
312. Hallmann, K., Zsurka, G., Moskau-Hartmann, S., Kirschner, J., Korinthenberg, R., Ruppert, A.K., Ozdemir, O., Weber, Y., Becker, F., Lerche, H., et al. (2014). A homozygous splice-site mutation in CARS2 is associated with progressive myoclonic epilepsy. Neurology 83, 2183-2187.
313. Taft, R.J., Vanderver, A., Leventer, R.J., Damiani, S.A., Simons, C., Grimmond, S.M., Miller, D., Schmidt, J., Lockhart, P.J., Pope, K., et al. (2013). Mutations in DARS cause hypomyelination with brain stem and spinal cord involvement and leg spasticity. Am J Hum Genet 92, 774-780.
314. Frohlich, D., Suchowerska, A.K., Voss, C., He, R., Wolvetang, E., von Jonquieres, G., Simons, C., Fath, T., Housley, G.D., and Klugmann, M. (2018). Expression Pattern of the Aspartyl-tRNA Synthetase DARS in the Human Brain. Front Mol Neurosci 11, 81.
315. Frohlich, D., Suchowerska, A.K., Spencer, Z.H., von Jonquieres, G., Klugmann, C.B., Bongers, A., Delerue, F., Stefen, H., Ittner, L.M., Fath, T., et al. (2017). In vivocharacterization of the aspartyl-tRNA synthetase DARS: Homing in on the leukodystrophy HBSL. Neurobiol Dis 97, 24-35.
316. Scheper, G.C., van der Klok, T., van Andel, R.J., van Berkel, C.G., Sissler, M., Smet, J., Muravina, T.I., Serkov, S.V., Uziel, G., Bugiani, M., et al. (2007). Mitochondrial aspartyl-tRNA synthetase deficiency causes leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation. Nature genetics 39, 534-539.
317. Yelam, A., Nagarajan, E., Chuquilin, M., and Govindarajan, R. (2019). Leucoencephalopathy with brain stem and spinal cord involvement and lactate elevation: a novel mutation in the DARS2 gene. BMJ Case Rep 12.
318. Shimojima, K., Higashiguchi, T., Kishimoto, K., Miyatake, S., Miyake, N., Takanashi, J.I., Matsumoto, N., and Yamamoto, T. (2017). A novel DARS2 mutation in a Japanese patient with leukoencephalopathy with brainstem and spinal cord involvement but no lactate elevation. Hum Genome Var 4, 17051.
319. Lan, M.Y., Chang, Y.Y., Yeh, T.H., Lin, T.K., and Lu, C.S. (2017). Leukoencephalopathy with brainstem and spinal cord involvement and lactate elevation (LBSL) with a novel DARS2 mutation and isolated progressive spastic paraparesis. J Neurol Sci 372, 229-231.
320. Kohler, C., Heyer, C., Hoffjan, S., Stemmler, S., Lucke, T., Thiels, C., Kohlschutter, A., Lobel, U., Horvath, R., Kleinle, S., et al. (2015). Early-onset leukoencephalopathy due to a homozygous missense mutation in the DARS2 gene. Mol Cell Probes 29, 319-322.
321. Yamashita, S., Miyake, N., Matsumoto, N., Osaka, H., Iai, M., Aida, N., and Tanaka, Y. (2013). Neuropathology of leukoencephalopathy with brainstem and spinal cord involvement and high lactate caused by a homozygous mutation of DARS2. Brain Dev 35, 312-316.
322. Miyake, N., Yamashita, S., Kurosawa, K., Miyatake, S., Tsurusaki, Y., Doi, H., Saitsu, H., and Matsumoto, N. (2011). A novel homozygous mutation of DARS2 may cause a severe LBSL variant. Clin Genet 80, 293-296.

323. Tzoulis, C., Tran, G.T., Gjerde, I.O., Aasly, J., Neckelmann, G., Rydland, J., Varga, V., Wadel-Andersen, P., and Bindoff, L.A. (2012). Leukoencephalopathy with brainstem and spinal cord involvement caused by a novel mutation in the DARS2 gene. J Neurol 259, 292-296.
324. Synofzik, M., Schicks, J., Lindig, T., Biskup, S., Schmidt, T., Hansel, J., Lehmann-Horn, F., and Schols, L. (2011). Acetazolamide-responsive exercise-induced episodic ataxia associated with a novel homozygous DARS2 mutation. J Med Genet 48, 713-715.
325. Lin, J., Chiconelli Faria, E., Da Rocha, A.J., Rodrigues Masruha, M., Pereira Vilanova, L.C., Scheper, G.C., and Van der Knaap, M.S. (2010). Leukoencephalopathy with brainstem and spinal cord involvement and normal lactate: a new mutation in the DARS2 gene. J Child Neurol 25, 1425-1428.
326. Mendes, M.I., Gutierrez Salazar, M., Guerrero, K., Thiffault, I., Salomons, G.S., Gauquelin, L., Tran, L.T., Forget, D., Gauthier, M.S., Waisfisz, Q., et al. (2018). Bi-allelic Mutations in EPRS, Encoding the Glutamyl-Prolyl-Aminoacyl-tRNA Synthetase, Cause a Hypomyelinating Leukodystrophy. Am J Hum Genet 102, 676-684.
327. Steenweg, M.E., Ghezzi, D., Haack, T., Abbink, T.E., Martinelli, D., van Berkel, C.G., Bley, A., Diogo, L., Grillo, E., Te Water Naude, J., et al. (2012). Leukoencephalopathy with thalamus and brainstem involvement and high lactate 'LTBL' caused by EARS2 mutations. Brain 135, 1387-1394.
328. Gungor, O., Ozkaya, A.K., Sahin, Y., Gungor, G., Dilber, C., and Aydin, K. (2016). A compound heterozygous EARS2 mutation associated with mild leukoencephalopathy with thalamus and brainstem involvement and high lactate (LTBL). Brain Dev 38, 857-861.
329. Taskin, B.D., Karalok, Z.S., Gurkas, E., Aydin, K., Aydogmus, U., Ceylaner, S., Karaer, K., Yilmaz, C., and Pearl, P.L. (2016). Early-Onset Mild Type Leukoencephalopathy Caused by a Homozygous EARS2 Mutation. J Child Neurol 31, 938-941.
330. Kevelam, S.H., Klouwer, F.C., Fock, J.M., Salomons, G.S., Bugiani, M., and van der Knaap, M.S. (2016). Absent Thalami Caused by a Homozygous EARS2 Mutation: Expanding Disease Spectrum of LTBL. Neuropediatrics 47, 64-67.
331. Biancheri, R., Lamantea, E., Severino, M., Diodato, D., Pedemonte, M., Cassandrini, D., Ploederl, A., Trucco, F., Fiorillo, C., Minetti, C., et al. (2015). Expanding the Clinical and Magnetic Resonance Spectrum of Leukoencephalopathy with Thalamus and Brainstem Involvement and High Lactate (LTBL) in a Patient Harboring a Novel EARS2 Mutation. JIMD Rep 23, 85-89.
332. Talim, B., Pyle, A., Griffin, H., Topaloglu, H., Tokatli, A., Keogh, M.J., Santibanez-Koref, M., Chinnery, P.F., and Horvath, R. (2013). Multisystem fatal infantile disease caused by a novel homozygous EARS2 mutation. Brain 136, e228.
333. Antonellis, A., Oprescu, S.N., Griffin, L.B., Heider, A., Amalfitano, A., and Innis, J.W. (2018). Compound heterozygosity for loss-of-function FARSB variants in a patient with classic features of recessive aminoacyl-tRNA synthetase-related disease. Human mutation 39, 834-840.
334. Zadjali, F., Al-Yahyaee, A., Al-Nabhani, M., Al-Mubaihsi, S., Gujjar, A., Raniga, S., and Al-Maawali, A. (2018). Homozygosity for FARSB mutation leads to Phe-tRNA synthetase-related disease of growth restriction, brain calcification, and interstitial lung disease. Human mutation 39, 1355-1359.
335. Elo, J.M., Yadavalli, S.S., Euro, L., Isohanni, P., Gotz, A., Carroll, C.J., Valanne, L., Alkuraya, F.S., Uusimaa, J., Paetau, A., et al. (2012). Mitochondrial phenylalanyl-tRNA synthetase mutations underlie fatal infantile Alpers encephalopathy. Hum Mol Genet 21, 4521-4529.
336. Yang, Y., Liu, W., Fang, Z., Shi, J., Che, F., He, C., Yao, L., Wang, E., and Wu, Y. (2016). A Newly Identified Missense Mutation in FARS2 Causes Autosomal-Recessive Spastic Paraplegia. Human mutation 37, 165-169.
337. Raviglione, F., Conte, G., Ghezzi, D., Parazzini, C., Righini, A., Vergaro, R., Legati, A., Spaccini, L., Gasperini, S., Garavaglia, B., et al. (2016). Clinical findings in a patient with FARS2 mutations and early-infantile-encephalopathy with epilepsy. Am J Med Genet A 170, 3004-3007.
338. Finsterer, J., Scorza, F.A., and Scorza, C.A. (2018). Antiepileptic treatment may determine the outcome of FARS2 mutation carriers. Mol Genet Metab Rep 17, 45.
339. Cho, J.S., Kim, S.H., Kim, H.Y., Chung, T., Kim, D., Jang, S., Lee, S.B., Yoo, S.K., Shin, J., Kim, J.I., et al. (2017). FARS2 mutation and epilepsy: Possible link with early-onset epileptic encephalopathy. Epilepsy Res 129, 118-124.
340. Almalki, A., Alston, C.L., Parker, A., Simonic, I., Mehta, S.G., He, L., Reza, M., Oliveira, J.M., Lightowlers, R.N., McFarland, R., et al. (2014). Mutation of the human mitochondrial phenylalanine-tRNA synthetase causes infantile-onset epilepsy and cytochrome c oxidase deficiency. Biochim Biophys Acta 1842, 56-64.
341. McMillan, H.J., Schwartzentruber, J., Smith, A., Lee, S., Chakraborty, P., Bulman, D.E., Beaulieu, C.L., Majewski, J., Boycott, K.M., and Geraghty, M.T. (2014). Compound heterozygous mutations in glycyl-tRNA synthetase are a proposed cause of systemic mitochondrial disease. BMC Med Genet 15, 36.
342. Antonellis, A., Ellsworth, R.E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S.Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L.T., et al. (2003). Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. Am J Hum Genet 72, 1293-1299.
343. Liao, Y.C., Liu, Y.T., Tsai, P.C., Chang, C.C., Huang, Y.H., Soong, B.W., and Lee, Y.C. (2015). Two Novel De Novo GARS Mutations Cause Early-Onset Axonal Charcot-Marie-Tooth Disease. PLoS One 10, e0133423.
344. Corcia, P., Brulard, C., Beltran, S., Marouillat, S., Bakkouche, S.E., Andres, C.R., Blasco, H., and Vourc'h, P. (2019). Typical bulbar ALS can be linked to GARS mutation. Amyotroph Lateral Scler Frontotemporal Degener, 1-3.
345. Rahane, C.S., Kutzner, A., and Heese, K. (2019). Establishing a human adrenocortical carcinoma (ACC)-specific gene mutation signature. Cancer Genet 230, 1-12.
346. Nan, H., Takaki, R., Hata, T., Ichinose, Y., Tsuchiya, M., Koh, K., and Takiyama, Y. (2018). Novel GARS mutation presenting as autosomal dominant intermediate Charcot-Marie-Tooth disease. J Peripher Nerv Syst.
347. Yu, X., Chen, B., Tang, H., Li, W., Fu, Y., Zhang, Z., and Yan, Y. (2018). A Novel Mutation of GARS in a Chinese Family With Distal Hereditary Motor Neuropathy Type V. Front Neurol 9, 571.

348. Holloway, M.P., DeNardo, B.D., Phornphutkul, C., Nguyen, K., Davis, C., Jackson, C., Richendrfer, H., Creton, R., and Altura, R.A. (2016). An asymptomatic mutation complicating severe chemotherapy-induced peripheral neuropathy (CIPN): a case for personalised medicine and a zebrafish model of CIPN. NPJ Genom Med 1, 16016.
349. Sun, A., Liu, X., Zheng, M., Sun, Q., Huang, Y., and Fan, D. (2015). A novel mutation of the glycyl-tRNA synthetase (GARS) gene associated with Charcot-Marie-Tooth type 2D in a Chinese family. Neurol Res 37, 782-787.
350. Kawakami, N., Komatsu, K., Yamashita, H., Uemura, K., Oka, N., Takashima, H., and Takahashi, R. (2014). [A novel mutation in glycyl-tRNA synthetase caused Charcot-Marie-Tooth disease type 2D with facial and respiratory muscle involvement]. Rinsho Shinkeigaku 54, 911-915.
351. Eskuri, J.M., Stanley, C.M., Moore, S.A., and Mathews, K.D. (2012). Infantile onset CMT2D/dSMA V in monozygotic twins due to a mutation in the anticodon-binding domain of GARS. J Peripher Nerv Syst 17, 132-134.
352. Hamaguchi, A., Ishida, C., Iwasa, K., Abe, A., and Yamada, M. (2010). Charcot-Marie-Tooth disease type 2D with a novel glycyl-tRNA synthetase gene (GARS) mutation. J Neurol 257, 1202-1204.
353. Achilli, F., Bros-Facer, V., Williams, H.P., Banks, G.T., AlQatari, M., Chia, R., Tucci, V., Groves, M., Nickols, C.D., Seburn, K.L., et al. (2009). An ENU-induced mutation in mouse glycyl-tRNA synthetase (GARS) causes peripheral sensory and motor phenotypes creating a model of Charcot-Marie-Tooth type 2D peripheral neuropathy. Dis Model Mech 2, 359-373.
354. Abe, A., Numakura, C., Saito, K., Koide, H., Oka, N., Honma, A., Kishikawa, Y., and Hayasaka, K. (2009). Neurofilament light chain polypeptide gene mutations in Charcot-Marie-Tooth disease: nonsense mutation probably causes a recessive phenotype. J Hum Genet 54, 94-97.
355. Dubourg, O., Azzedine, H., Yaou, R.B., Pouget, J., Barois, A., Meininger, V., Bouteiller, D., Ruberg, M., Brice, A., and LeGuern, E. (2006). The G526R glycyl-tRNA synthetase gene mutation in distal hereditary motor neuropathy type V. Neurology 66, 1721-1726.
356. Del Bo, R., Locatelli, F., Corti, S., Scarlato, M., Ghezzi, S., Prelle, A., Fagiolari, G., Moggio, M., Carpo, M., Bresolin, N., et al. (2006). Coexistence of CMT-2D and distal SMA-V phenotypes in an Italian family with a GARS gene mutation. Neurology 66, 752-754.
357. Vester, A., Velez-Ruiz, G., McLaughlin, H.M., Program, N.C.S., Lupski, J.R., Talbot, K., Vance, J.M., Zuchner, S., Roda, R.H., Fischbeck, K.H., et al. (2013). A loss-of-function variant in the human histidyl-tRNA synthetase (HARS) gene is neurotoxic in vivo. Hum Mutat 34, 191-199.
358. Puffenberger, E.G., Jinks, R.N., Sougnez, C., Cibulskis, K., Willert, R.A., Achilly, N.P., Cassidy, R.P., Fiorentini, C.J., Heiken, K.F., Lawrence, J.J., et al. (2012). Genetic mapping and exome sequencing identify variants associated with five novel diseases. PLoS One 7, e28936.
359. Pierce, S.B., Chisholm, K.M., Lynch, E.D., Lee, M.K., Walsh, T., Opitz, J.M., Li, W., Klevit, R.E., and King, M.C. (2011). Mutations in mitochondrial histidyl tRNA synthetase HARS2 cause ovarian dysgenesis and sensorineural hearing loss of Perrault syndrome. Proc Natl Acad Sci U S A 108, 6543-6548.
360. Demain, L.A.M., Gerkes, E.H., Smith, R.J.H., Molina-Ramirez, L.P., O'Keefe, R.T., and Newman, W.G. (2020). A recurrent missense variant in HARS2 results in variable sensorineural hearing loss in three unrelated families. J Hum Genet 65, 305-311.
361. Kopajtich, R., Murayama, K., Janecke, A.R., Haack, T.B., Breuer, M., Knisely, A.S., Harting, I., Ohashi, T., Okazaki, Y., Watanabe, D., et al. (2016). Biallelic IARS Mutations Cause Growth Retardation with Prenatal Onset, Intellectual Disability, Muscular Hypotonia, and Infantile Hepatopathy. Am J Hum Genet 99, 414-422.
362. Schwartzentruber, J., Buhas, D., Majewski, J., Sasarman, F., Papillon-Cavanagh, S., Thiffault, I., Sheldon, K.M., Massicotte, C., Patry, L., Simon, M., et al. (2014). Mutation in the nuclear-encoded mitochondrial isoleucyl-tRNA synthetase IARS2 in patients with cataracts, growth hormone deficiency with short stature, partial sensorineural deafness, and peripheral neuropathy or with Leigh syndrome. Hum Mutat 35, 1285-1289.
363. Moosa, S., Haagerup, A., Gregersen, P.A., Petersen, K.K., Altmuller, J., Thiele, H., Nurnberg, P., Cho, T.J., Kim, O.H., Nishimura, G., et al. (2017). Confirmation of CAGSSS syndrome as a distinct entity in a Danish patient with a novel homozygous mutation in IARS2. Am J Med Genet A 173, 1102-1108.
364. Jabbour, S., and Harissi-Dagher, M. (2016). Recessive Mutation in a Nuclear-Encoded Mitochondrial tRNA Synthetase Associated With Infantile Cataract, Congenital Neurotrophic Keratitis, and Orbital Myopathy. Cornea 35, 894-896.
365. McLaughlin, H.M., Sakaguchi, R., Liu, C., Igarashi, T., Pehlivan, D., Chu, K., Iyer, R., Cruz, P., Cherukuri, P.F., Hansen, N.F., et al. (2010). Compound heterozygosity for loss-of-function lysyl-tRNA synthetase mutations in a patient with peripheral neuropathy. Am J Hum Genet 87, 560-566.
366. McMillan, H.J., Humphreys, P., Smith, A., Schwartzentruber, J., Chakraborty, P., Bulman, D.E., Beaulieu, C.L., Consortium, F.C., Majewski, J., Boycott, K.M., et al. (2015). Congenital Visual Impairment and Progressive Microcephaly Due to Lysyl-Transfer Ribonucleic Acid (RNA) Synthetase (KARS) Mutations: The Expanding Phenotype of Aminoacyl-Transfer RNA Synthetase Mutations in Human Disease. J Child Neurol 30, 1037-1043.
367. Santos-Cortez, R.L., Lee, K., Azeem, Z., Antonellis, P.J., Pollock, L.M., Khan, S., Irfanullah, Andrade-Elizondo, P.B., Chiu, I., Adams, M.D., et al. (2013). Mutations in KARS, encoding lysyl-tRNA synthetase, cause autosomal-recessive nonsyndromic hearing impairment DFNB89. Am J Hum Genet 93, 132-140.
368. Itoh, M., Dai, H., Horike, S.I., Gonzalez, J., Kitami, Y., Meguro-Horike, M., Kuki, I., Shimakawa, S., Yoshinaga, H., Ota, Y., et al. (2019). Biallelic KARS pathogenic variants cause an early-onset progressive leukodystrophy. Brain 142, 560-573.
369. Casey, J.P., McGettigan, P., Lynam-Lennon, N., McDermott, M., Regan, R., Conroy, J., Bourke, B., O'Sullivan, J., Crushell, E., Lynch, S., et al. (2012). Identification of a mutation in LARS as a novel cause of infantile hepatopathy. Mol Genet Metab 106, 351-358.
370. Pierce, S.B., Gersak, K., Michaelson-Cohen, R., Walsh, T., Lee, M.K., Malach, D., Klevit, R.E., King, M.C., and Levy-Lahad, E. (2013). Mutations in LARS2, encoding mitochondrial leucyl-tRNA synthetase, lead to premature ovarian failure and hearing loss in Perrault syndrome. Am J Hum Genet 92, 614-620.

371. Gonzalez, M., McLaughlin, H., Houlden, H., Guo, M., Yo-Tsen, L., Hadjivassilious, M., Speziani, F., Yang, X.L., Antonellis, A., Reilly, M.M., et al. (2013). Exome sequencing identifies a significant variant in methionyl-tRNA synthetase (MARS) in a family with late-onset CMT2. J Neurol Neurosurg Psychiatry 84, 1247-1249.
372. van Meel, E., Wegner, D.J., Cliften, P., Willing, M.C., White, F.V., Kornfeld, S., and Cole, F.S. (2013). Rare recessive loss-of-function methionyl-tRNA synthetase mutations presenting as a multi-organ phenotype. BMC Med Genet 14, 106.
373. Hadchouel, A., Wieland, T., Griese, M., Baruffini, E., Lorenz-Depiereux, B., Enaud, L., Graf, E., Dubus, J.C., Halioui-Louhaichi, S., Coulomb, A., et al. (2015). Biallelic Mutations of Methionyl-tRNA Synthetase Cause a Specific Type of Pulmonary Alveolar Proteinosis Prevalent on Reunion Island. Am J Hum Genet 96, 826-831.
374. Alzaid, M., Alshamrani, A., Al Harbi, A.S., Alenzi, A., and Mohamed, S. (2019). Methionyl-tRNA synthetase novel mutation causes pulmonary alveolar proteinosis. Saudi Med J 40, 195-198.
375. Bayat, V., Thiffault, I., Jaiswal, M., Tetreault, M., Donti, T., Sasarman, F., Bernard, G., Demers-Lamarche, J., Dicaire, M.J., Mathieu, J., et al. (2012). Mutations in the mitochondrial methionyl-tRNA synthetase cause a neurodegenerative phenotype in flies and a recessive ataxia (ARSAL) in humans. PLoS Biol 10, e1001288.
376. Webb, B.D., Wheeler, P.G., Hagen, J.J., Cohen, N., Linderman, M.D., Diaz, G.A., Naidich, T.P., Rodenburg, R.J., Houten, S.M., and Schadt, E.E. (2015). Novel, compound heterozygous, single-nucleotide variants in MARS2 associated with developmental delay, poor growth, and sensorineural hearing loss. Hum Mutat 36, 587-592.
377. Finsterer, J. (2018). Management of NARS2-Related Mitochondrial Disorder is Complex. Pediatr Neurol.
378. Seaver, L.H., DeRoos, S., Andersen, N.J., Betz, B., Prokop, J., Lannen, N., Jordan, R., and Rajasekaran, S. (2018). Lethal NARS2-Related Disorder Associated With Rapidly Progressive Intractable Epilepsy and Global Brain Atrophy. Pediatr Neurol 89, 26-30.
379. Mizuguchi, T., Nakashima, M., Kato, M., Yamada, K., Okanishi, T., Ekhilevitch, N., Mandel, H., Eran, A., Toyono, M., Sawaishi, Y., et al. (2017). PARS2 and NARS2 mutations in infantile-onset neurodegenerative disorder. J Hum Genet 62, 525-529.
380. Zhang, X., Ling, J., Barcia, G., Jing, L., Wu, J., Barry, B.J., Mochida, G.H., Hill, R.S., Weimer, J.M., Stein, Q., et al. (2014). Mutations in QARS, encoding glutaminyl-tRNA synthetase, cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures. Am J Hum Genet 94, 547-558.
381. Nafisinia, M., Sobreira, N., Riley, L., Gold, W., Uhlenberg, B., Weiss, C., Boehm, C., Prelog, K., Ouvrier, R., and Christodoulou, J. (2017). Mutations in RARS cause a hypomyelination disorder akin to Pelizaeus-Merzbacher disease. Eur J Hum Genet 25, 1134-1141.
382. Mendes, M.I., Green, L.M.C., Bertini, E., Tonduti, D., Aiello, C., Smith, D., Salsano, E., Beerepoot, S., Hertecant, J., von Spiczak, S., et al. (2020). RARS1-related hypomyelinating leukodystrophy: Expanding the spectrum. Annals of clinical and translational neurology 7, 83-93.
383. Matsumoto, N., Watanabe, N., Iibe, N., Tatsumi, Y., Hattori, K., Takeuchi, Y., Oizumi, H., Ohbuchi, K., Torii, T., Miyamoto, Y., et al. (2019). Hypomyelinating leukodystrophy-associated mutation of RARS leads it to the lysosome, inhibiting oligodendroglial morphological differentiation. Biochem Biophys Rep 20, 100705.
384. Edvardson, S., Shaag, A., Kolesnikova, O., Gomori, J.M., Tarassov, I., Einbinder, T., Saada, A., and Elpeleg, O. (2007). Deleterious mutation in the mitochondrial arginyl-transfer RNA synthetase gene is associated with pontocerebellar hypoplasia. Am J Hum Genet 81, 857-862.
385. Luhl, S., Bode, H., Schlotzer, W., Bartsakoulia, M., Horvath, R., Abicht, A., Stenzel, M., Kirschner, J., and Grunert, S.C. (2016). Novel homozygous RARS2 mutation in two siblings without pontocerebellar hypoplasia - further expansion of the phenotypic spectrum. Orphanet J Rare Dis 11, 140.
386. Li, Z., Schonberg, R., Guidugli, L., Johnson, A.K., Arnovitz, S., Yang, S., Scafidi, J., Summar, M.L., Vezina, G., Das, S., et al. (2015). A novel mutation in the promoter of RARS2 causes pontocerebellar hypoplasia in two siblings. J Hum Genet 60, 363-369.
387. Musante, L., Puttmann, L., Kahrizi, K., Garshasbi, M., Hu, H., Stehr, H., Lipkowitz, B., Otto, S., Jensen, L.R., Tzschach, A., et al. (2017). Mutations of the aminoacyl-tRNA-synthetases SARS and WARS2 are implicated in the etiology of autosomal recessive intellectual disability. Hum Mutat 38, 621-636.
388. Belostotsky, R., Ben-Shalom, E., Rinat, C., Becker-Cohen, R., Feinstein, S., Zeligson, S., Segel, R., Elpeleg, O., Nassar, S., and Frishberg, Y. (2011). Mutations in the mitochondrial seryl-tRNA synthetase cause hyperuricemia, pulmonary hypertension, renal failure in infancy and alkalosis, HUPRA syndrome. Am J Hum Genet 88, 193-200.
389. Diodato, D., Melchionda, L., Haack, T.B., Dallabona, C., Baruffini, E., Donnini, C., Granata, T., Ragona, F., Balestri, P., Margollicci, M., et al. (2014). VARS2 and TARS2 mutations in patients with mitochondrial encephalomyopathies. Hum Mutat 35, 983-989.
390. Wang, Y., Zhou, X.L., Ruan, Z.R., Liu, R.J., Eriani, G., and Wang, E.D. (2016). A Human Disease-causing Point Mutation in Mitochondrial Threonyl-tRNA Synthetase Induces Both Structural and Functional Defects. J Biol Chem 291, 6507-6520.
391. Taylor, R.W., Pyle, A., Griffin, H., Blakely, E.L., Duff, J., He, L., Smertenko, T., Alston, C.L., Neeve, V.C., Best, A., et al. (2014). Use of whole-exome sequencing to determine the genetic basis of multiple mitochondrial respiratory chain complex deficiencies. JAMA 312, 68-77.
392. Nowaczyk, M.J., Huang, L., Tarnopolsky, M., Schwartzentruber, J., Majewski, J., Bulman, D.E., Forge Canada Consortium, C.R.C.C., Hartley, T., and Boycott, K.M. (2017). A novel multisystem disease associated with recessive mutations in the tyrosyl-tRNA synthetase (YARS) gene. Am J Med Genet A 173, 126-134.
393. Ma, K., Xie, M., He, X., Liu, G., Lu, X., Peng, Q., Zhong, B., and Li, N. (2018). A novel compound heterozygous mutation in VARS2 in a newborn with mitochondrial cardiomyopathy: a case report of a Chinese family. BMC Med Genet 19, 202.
394. Alsemari, A., Al-Younes, B., Goljan, E., Jaroudi, D., BinHumaid, F., Meyer, B.F., Arold, S.T., and Monies, D. (2017). Recessive VARS2 mutation underlies a novel syndrome with epilepsy, mental retardation, short stature, growth hormone deficiency, and hypogonadism. Hum Genomics 11, 28.

395. Tsai, P.C., Soong, B.W., Mademan, I., Huang, Y.H., Liu, C.R., Hsiao, C.T., Wu, H.T., Liu, T.T., Liu, Y.T., Tseng, Y.T., et al. (2017). A recurrent WARS mutation is a novel cause of autosomal dominant distal hereditary motor neuropathy. Brain 140, 1252-1266.
396. Theisen, B.E., Rumyantseva, A., Cohen, J.S., Alcaraz, W.A., Shinde, D.N., Tang, S., Srivastava, S., Pevsner, J., Trifunovic, A., and Fatemi, A. (2017). Deficiency of WARS2, encoding mitochondrial tryptophanyl tRNA synthetase, causes severe infantile onset leukoencephalopathy. Am J Med Genet A 173, 2505-2510.
397. Wortmann, S.B., Timal, S., Venselaar, H., Wintjes, L.T., Kopajtich, R., Feichtinger, R.G., Onnekink, C., Muhlmeister, M., Brandt, U., Smeitink, J.A., et al. (2017). Biallelic variants in WARS2 encoding mitochondrial tryptophanyl-tRNA synthase in six individuals with mitochondrial encephalopathy. Hum Mutat 38, 1786-1795.
398. Hubers, A., Huppertz, H.J., Wortmann, S.B., and Kassubek, J. (2020). Mutation of the WARS2 Gene as the Cause of a Severe Hyperkinetic Movement Disorder. Mov Disord Clin Pract 7, 88-90.
399. Nakajima, J., Eminoglu, T.F., Vatansever, G., Nakashima, M., Tsurusaki, Y., Saitsu, H., Kawashima, H., Matsumoto, N., and Miyake, N. (2014). A novel homozygous YARS2 mutation causes severe myopathy, lactic acidosis, and sideroblastic anemia 2. J Hum Genet 59, 229-232.
400. Sasarman, F., Nishimura, T., Thiffault, I., and Shoubridge, E.A. (2012). A novel mutation in YARS2 causes myopathy with lactic acidosis and sideroblastic anemia. Hum Mutat 33, 1201-1206.
401. Riley, L.G., Cooper, S., Hickey, P., Rudinger-Thirion, J., McKenzie, M., Compton, A., Lim, S.C., Thorburn, D., Ryan, M.T., Giege, R., et al. (2010). Mutation of the mitochondrial tyrosyl-tRNA synthetase gene, YARS2, causes myopathy, lactic acidosis, and sideroblastic anemia--MLASA syndrome. Am J Hum Genet 87, 52-59.