Delineating the clinical, genetic and molecular characteristics of neurodevelopmental disorders in a community setting

Submitted by Lettie Eleanor Rawlins

to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Medical Studies

In May 2022

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signature: ……………………………………………………………..

ACKNOWLEDGEMENTS

time Neuroscience friend Dr Rachel Darby who was extraordinarily kind (and crazy) enough to proofread this for me.

Lastly, I would like to thank my grandparents, Valerie and Walter, who taught me to believe in myself and that you can achieve most things if you work hard for them, may you rest in peace.

I dedicate this thesis to my children,

Josephine and Thomas

ABSTRACT

Inherited neurodevelopmental disorders are a large group of clinically and genetically heterogeneous conditions affecting the development, structure and functioning of the central nervous system, and are frequently associated with extremes of brain growth. Whilst specific neurodevelopmental disorders are typically individually rare, as a group these conditions comprise a significant healthcare burden globally. A significant contribution to our understanding of neurodevelopmental disorders has been made by the study of autosomal recessive conditions, which are rare in the general population but occur at increased frequency in certain genetically isolated communities due to founder gene variants. The studies described in this thesis entail clinical, genetic and functional studies of inherited single gene disorders of brain growth identified in North American Amish and rural Pakistani communities, leading to investigations in other families globally.

Chapter three describes the identification of a microcephalic neurodevelopmental disorder resulting from biallelic variants in the *TRAPPC10* gene, encoding a transport protein particle (TRAPP) complex subunit involved in membrane trafficking and other cellular processes. This study, stemming from genetic findings in two Pakistani families harbouring distinct *TRAPPC10* variants, entailed comprehensive clinical, genomic, mouse and functional studies to confirm *TRAPPC10*-related disorder as a novel TRAPPopathy. This chapter also describes preliminary clinical and genetic studies identifying or consolidating five further TRAPPopathy disorders.

KPTN-related megalencephalic neurodevelopmental disorder was first identified by our research group in 2014. Chapter four documents the most detailed phenotypic description of the condition to date, alongside genetic findings of 36 affected individuals identified through international collaboration. This work enabled preliminary genotype-phenotype correlations to be drawn and patient management guidelines to be developed. Additionally, detailed comparisons with a *Kptn*$^{-/-}$ knock-out mouse model that closely recapitulates the human disease, alongside cell studies, provided fundamental new insights into the condition. Together, this enabled *KPTN*-related disorder to be confirmed as a new

mTORopathy, highlighting the potential of mTOR inhibitors as a candidate therapeutic option for the condition.

Chapter five details extensive clinical and genetic datasets to consolidate and more comprehensively define two previously poorly described causes of autosomal recessive disorders of brain growth and development. These involve variants in *CEP55* associated with a severe lethal fetal disorder characterised by a variable spectrum of brain and kidney abnormalities, and *INPP4A* associated with a variable neurodevelopmental disorder with features of severe cognitive impairment, seizures, muscle weakness, cerebellar signs, spasticity and behavioural abnormalities.

The clinical, genetic and molecular delineation of rare neurodevelopmental disorders facilitates accurate diagnosis and management of these conditions. The work described in this thesis also provides invaluable insights into disease pathogenesis and improves the scientific understanding of the subcellular processes underlying neurodevelopment. This knowledge is crucial to ultimately enable the discovery and delivery of potential new therapeutic avenues to treat these disorders.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

<	Less than
≤	Less than or equal to
>	Greater than
≥	Greater than or equal to
=	Equals
~	Approximately
°C	Degrees Celsius
%	Percent
#	number
µCT	Microcomputed tomography
µg	micrograms
µl	microlitres
µm	micrometres
µM	micromolar
3'	Three prime
5'	Five prime
AC	Allele count
ACD	Acid Citrate Dextrose
ACGS	Association for Clinical Genomic Science
ACMG	American College of Medical Genetics and Genomics
AD	Autosomal dominant
ADHD	Attention deficit hyperactivity disorder
AF	Allele frequency
AKT	A serine/threonine protein kinase (protein kinase B)
Ala	Alanine
AMPK	AMP-activated protein kinase
APS	Ammonium per sulphate
AR	Autosomal recessive
Arg	Arginine
ASH	ASPM, SPD-2, Hydin domain
Asn	Asparagine
AVM	Ateriovenous malformation
BAM	Binary alignment map
BBS	Bardet-Biedel syndrome
BLAT	BLAST-like alignment tool
BMI	Body mass index
bp	Base pair
BP	attB to attP reaction
BSA	Bovine serum albumin
BWA	Burrows-Wheeler alignment
C	Carboxy group
CA	California
CAP	Complementary analysis project
Cas9	CRISPR-associated protein 9
CCD	Charged-coupled device
cDNA	Complementary DNA
CGH	Comparative genomic hybridisation
chr	Chromosome
CI	Confidence interval

CIP	Calf intestinal alkaline phosphatase
CK	Creatine kinase
cm	Centimetres
CMAP	Connectivity Map
CMV	Cytomegalovirus
CNS	Central nervous system
CNV	Copy number variant
CO_2	Carbon dioxide
COG	Conserved oligomeric Golgi
COP	Coat proteins
CP	Cortical plate
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSF	Cerebrospinal fluid
CT	Computed tomography
Cys	Cysteine
dATP	Deoxyadenosine triphosphate
dbSNP	Single Nucleotide Polymorphism Database
dCTP	Deoxycytidine triphosphate
DDD	Deciphering developmental disorders study
ddH_2O	Double distilled water
del	Deletion
dGTP	Deoxyguanosine triphosphate
DM	Disease causing mutation
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DP	Depth
DSM-V	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
e.g.	For example
EBV	Epstein Barr virus
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EEG	Electroencephalogram
ER	Endoplasmic reticulum
ERG	Electroretinogram
ERK	Extracellular signal-regulated kinase
ESCRT	Endosomal sorting complex required for transport
Exo I	Exonuclease I
F	Female
FBS	Fetal bovine serum
FLAG	DYKDDDDK-tag
fs	Frameshift
FSIQ	Full-scale intelligence quotient
g	gram
GAP	GTPase activating protein
GATK	Genome Analysis Toolkit
GATOR 1	Gap Activity TOward Rags 1
GC	Guanine-cytosine

GDD	Global developmental delay
GDPR	General Data Protection Regulation
GE	Genomics England
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
gnomAD	Genome Aggregation Database
GO	Gene ontology
GRCh37/38	Genome Reference Consortium human genome build 37/38
gRNA	Guide RNA
GTEx	Genotype-Tissue Expression
GTP	Guanosine-5'-triphosphate
GTCS	Generalised tonic-clonic seizures
H_2O	Water
HAP1	Haploid 1 cells
HDL	High-density lipoprotein
HEK293	Human embryonic kidney 293
HEMI	Hemizygous
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HET	Heterozygous
hg38	Human genome build 38
HGMD	Human Gene Mutation Database
HGNC	HUGO Gene Nomenclature Committee
His	Histidine
HIV	Human immunodeficiency virus
HOM	Homozygous
HSF	Human Splice Finder
HTA	Human tissue authority
Hz	Hertz
ICD-11	International Statistical Classification of Diseases and Related Health Problems, 11th edition
ID	Identification
IDT	Integrated DNA Technologies
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IGV	Integrative Genome Viewer
IKKβ	Inhibitor of nuclear factor kappa-B kinase
Ile	Isoleucine
IMDM	Iscove's modification of Dulbecco's modified eagle medium
Indel	Insertion and deletion
iPSC	Induced pluripotent stem cell
IQ	Intelligence quotient
iSVZ	Inner subventricular zone
IUFD	Intrauterine fetal death
IUGR	Intrauterine growth restriction
IZ	Intermediate zone
JAK	Janus kinase
kb	kilobase
KCl	Potassium chloride

kDa	kilodalton
kg	kilogram
KH_2PO_4	Monopotassium phosphate
KICSTOR	KPTN, ITFG2, C12orf66, and SZT2 target of rapamycin
KO	Knockout
KPK	Khyber Pakhtunkhwa
KPTN	Kaptin
L	Litre
LAB	Lithium acetate borate
lb	Pounds
LB	Luria Bertani
LCL	Lymphoblastoid cell line
LDL	Low-density lipoprotein
Leu	Leucine
LMBP	Laboratory of Molecular Biology Plasmid
LMS	Lambda for the skew, Mu for the median, and Sigma for the generalized coefficient of variation
LOEUF	Loss-of-function observed/expected upper bound fraction
LOF	Loss of function
LR	attL to attR reaction
Lys	Lysine
m	male
M	Molar
MAF	Minor allele frequency
MAPK	Mitogen activated protein kinase
MARCH	Multinucleated neurons, anhydramnios, renal dysplasia, cerebellar hypoplasia and hydranencephaly
MASD	Macrocephaly, autistic features, seizures, developmental delay
Mat	Maternal
Mb	Megabase
MCPH	Microcephaly Primary Hereditary
Met	Methionine
mg	Milligrams
$MgCl_2$	Magnesium chloride
MIM	Mendelian inheritance in man
min	Minute
MKS	Meckel syndrome
ml	Millilitres
mm	Millimetres
mM	Millimolar
MOPD	Microcephalic osteodysplastic primordial dwarfism
MORM	Mental retardation, truncal obesity, congenital nonprogressive retinal dystrophy, and micropenis
MPPH	Megalencephaly polymicrogyria polydactyly hydrocephalus
MQ	Mapping quality
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
MSUD	Maple syrup urine disease
mTOR	Mammalian target of rapamycin
mTORC1	Mammalian target of rapamycin complex 1
mTORC2	Mammalian target of rapamycin complex 2

MZ	Marginal zone
N	Amine group
NA	Not available
Na$_2$HPO$_4$	Disodium hydrogen phosphate
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
ND	Neonatal death
NE	Neuroepithelium
NEFAC	Non-esterified free fatty acids
NEPSY	NEuroPSYchological Assessment
NF	Neurofibromatosis
NFE	Non-Finnish European
NFκβ	Nuclear factor kappa B
ng	Nanogram
NGS	Next generation sequencing
NIH	National Institute of Health
NIK	NFκB-inducing kinase
nIPC	Neuronal intermediate progenitor cell
NK	Not known
nm	Nanometres
NMD	Nonsense mediated mRNA decay
NMDA	N-methyl-D-aspartate
NMDAR	N-methyl-D-aspartate receptor
NP	Not performed
NPC	Neural precursor cells
OCD	Obsessive compulsive disorder
OD	Optical density
OFC	Occipitofrontal circumference
OMIM	Online Mendelian Inheritance in Man
ORF	Open reading frame
oSVZ	Outer subventricular zone
oz	ounces
p	Probability value
P	Postnatal day
PA	Propionic acidaemia
PacBio	Pacific Biosciences
PAGE	Polyacrylamide gel electrophoresis
Pat	Paternal
PBS	Phosphate buffer solution
PBST	Phosphate buffer solution with tween
PCR	Polymerase chain reaction
PD	Pairwise discrimination
PDK1	Phosphoinositide-dependent kinase
PF	Palpebral fissures
pg	picagrams
PhD	Doctor of Philosophy
Phe	Phenylalanine
PI	Phosphoinositide
PI3K	Phosphatidylinositol 3-kinase
PKU	Phenylketonuria
pLI	Probability of being loss-of-function intolerant

PM	Post-mortem
PMSE	Polyhydramnios, megalencephaly, and symptomatic epilepsy
PRI	Perceptual reasoning index
Pro	Proline
PROVEAN	Protein variation effect analyser
PS6	Phosphorylated ribosomal protein S6
PSI	Processing speed index
PtdIns	Phosphatidylinositol
PTEN	Phosphatase and Tensin Homolog
Ras	Rat sarcoma virus
RBANS	Repeatable Battery for Assessment of Neurological Status
REVEL	Rare Exome Variant Ensemble Learner
RFP	Red fluorescent protein
RGC	Radial glial cell
RIN	RNA integrity number
RNA	Ribonucleic acid
RP	Retinitis pigmentosa
RPE	Retinal pigment epithelium
rpm	Revolutions per minute
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
SAGA	Swiss Anabaptist Geneological Association
SAP	Shrimp alkaline phosphatase
SB	Stillbirth
SC	Synthetic complete
SCID	Severe combined immunodeficiency
SD	Standard deviation
SDS	Standard deviation score
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	seconds
SEDT	Spondyloepiphyseal dysplasia tarda
SEM	Standard error of the mean
Ser	Serine
SHH	Sonic hedgehog
SIFT	Sorting Intolerant from Tolerant
siRNA	Small interfering RNA
SNHL	Sensorineural hearing loss
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SoR	Social recognition
STAT	Signal transducer and activator of transcription
SUDEP	Sudden unexplained death in epilepsy
Ta	Annealing temperature
TBE	Tris borate EDTA
TE	Tris-EDTA
TEMED	N,N,NN-tetramethylethylenediamine
Thr	Threonine
Tm	Melting temperature
TM	Trademark
TOP	Termination of pregnancy
TPR	Tetratricopeptide repeats

TRAPP	Transport protein particle
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
Trp	Tryptophan
TSC	Tuberose sclerosis complex
Tyr	Tyrosine
U	Units
UCSC	University of California, Santa Cruz
UK	United Kingdom
UPD	Uniparental disomy
USA	United States of America
USS	Ultrasound scan
UTR	Untranslated region
UV	Ultraviolet
v	Version
V	volts
Val	Valine
VCF	Variant call format
VCI	Verbal comprehension index
VEP	Visual evoked potential
VSVG	Vesicular stomatitis virus G
VUS	Variant of uncertain significance
VZ	Ventricular zone
W	watt
WAIS	Weschler Adult Intelligence Scale
WES	Whole exome sequencing
WGS	Whole genome sequencing
WHO	World Health Organisation
WISC	Weschler Intelligence Scale for Children
WOH	Windows of Hope
WMI	Working memory index
WT	Wild type
X	Times
XLR	X-linked recessive
y	Years

1 INTRODUCTION

1.1 Brain development

Gastrulation occurs during the third week of human embryonic development; epiblast cells migrate towards the primitive streak and their invagination creates three germ layers: the ectoderm, mesoderm and endoderm that develop into all the organs and tissues of the embryo [1]. The neural plate arises from thickening of the dorsal ectoderm, which consists of neuroepithelial stem cells that form the neural tube and subsequently the nervous system through a process called neurulation [2] (Figure 1.1). Fusion starts at the caudal end of the hindbrain and advances cranially and caudally, with closure of the anterior (cephalic) and posterior (caudal) neuropores at day 25 and 28 respectively (Figure 1.1) [2]. During this process, lateral neuroectodermal cells pinch off from the neural fold to become the neural crest cells that form the peripheral nervous system. The neural tube then comprises a closed tubular structure with the caudal narrow end developing into the spinal cord and the cephalic end developing into the brain [2]. Neuroepithelial cells (neuroblasts) within the caudal neural tube divide forming the mantle layer that later becomes the grey matter of the spinal cord, with arising nerve fibres forming the outermost marginal layer which becomes the white matter [1]. The cephalic neural tube forms three primary brain vesicles (Table 1.1) and two flexures, in the midbrain region (cephalic flexure) and between the hindbrain and spinal cord (cervical flexure), which further develop by five weeks into five secondary vesicles (Table 1.1) that become distinct brain structures containing the ventricles and fluid filled cavities continuous with the central spinal canal [3].
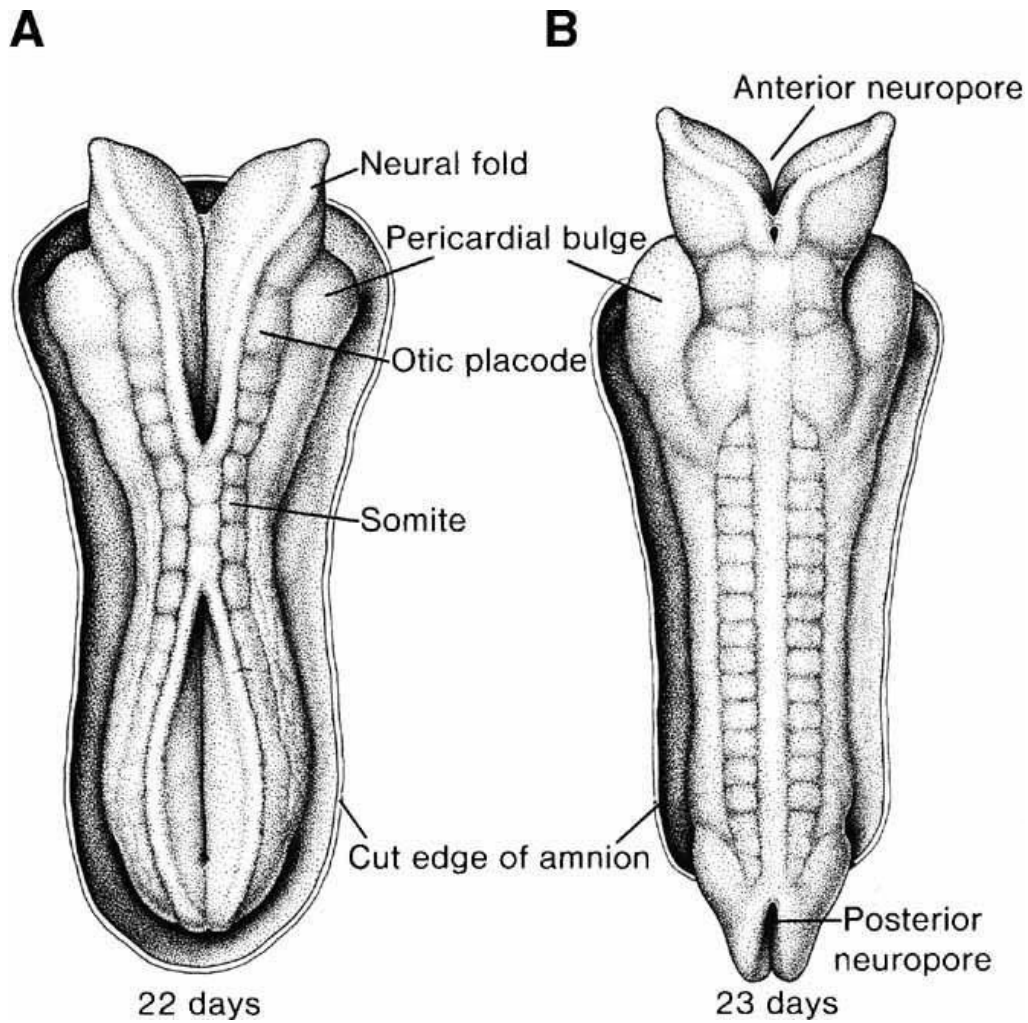
Figure 1.1: Dorsal views of the developing embryo demonstrating the formation and closure of the neural tube

Development of the neural tube from the embryonic ectoderm between the 3rd and 4th week of embryonic development. (**A**) 22 days: the mid-region of the neural plate becomes depressed to form the neural groove and elevation of the lateral edges form the neural folds which gradually move towards each other and fuse in the midline. (**B**) 23 days: the neural folds fuse in both directions to close the neural tube. Reproduced with permission from T.W. Sadler [2] (licence granted by John Wiley and Sons, Appendix 7).

Primary Vesicles	Secondary vesicles	Brain structures	Cavity derivatives
Prosencephalon (forebrain)	Telencephalon	Cerebral hemispheres and globus pallidus	Lateral ventricles
	Diencephalon	Thalamus, hypothalamus, epithalamus, optic cup and stalk	3rd ventricle
Mesencephalon (midbrain)	Mesencephalon	Midbrain	Cerebral aqueduct
Rhombencephalon (hindbrain)	Metencephalon	Pons and cerebellum	Upper part of the 4th ventricle
	Myelencephalon	Medulla	Lower part of the 4th ventricle

Table 1.1: Development of the structures of the brain and ventricles from the primary vesicles of the cephalic neural tube [1]

1.1.1 Development of the hindbrain

The myelencephalon is the most caudal brain region, which develops into the medulla (Table 1.1) [3]. The sulcus limitans divides the basal and alar plates containing the motor and sensory relay nuclei. The roof plate consists of ependymal cells that form the choroid plexus of the 4th ventricle that produces cerebrospinal fluid (CSF). The medulla contains ascending and descending nerve tracts transmitting signals between higher levels of the brain and the rest of the body and controls several autonomic functions such as vasodilation.

The metencephalon develops into the pons and cerebellum and similarly has basal and alar plates that develop into motor and sensory nuclei. The pons contains the pathway for nerve fibres between the cerebellar and cerebral cortices and the spinal cord. The cerebellum develops from the cerebellar plate; at 12 weeks this forms two hemispheres joined by the central midline vermis, with a small anterior lobe, separated by the primary fissure from the large posterior lobe and the transverse fissure separating the primitive flocculonodular lobe [4]. Neuroblasts migrate to the cerebellar surface to form the external granular layer, a proliferative zone of dividing cells that forms the cerebellar cortex grey matter consisting of three layers: the outer molecular layer, middle Purkinje cell layer and inner granular layer that surround the inner white matter [3]. The cerebellum

attaches to the dorsal brainstem through the inferior, middle and superior cerebellar peduncles, connecting the medulla, pons and midbrain respectively. The cerebellar cortex is highly convoluted with transverse narrow parallel folia and the cerebellar white matter consists of cortical nerve fibres and four paired cerebellar nuclei (dentate, emboliform, globose and fastigial) consisting of projecting motor neurons. The function of the cerebellum is purely motor, it unconsciously controls and maintains balance, coordinates movement and influences posture and muscle tone [4].

1.1.2 Development of the midbrain

The relatively small mesencephalon is located rostral to the pons. The ventral basal plates consist of two motor nuclei with the marginal layer developing into the crus cerebri that form nerve fibre pathways from the cerebral cortex to the brainstem and spinal cord. The dorsal alar plates develop into the superior and inferior colliculi, involved in visual and auditory reflexes respectively [3].

1.1.3 Development of the forebrain

The diencephalon develops from the median prosencephalon and forms the optic cup and stalk, thalamus, hypothalamus, epiphysis and pituitary gland (Table 1.1) [3]. The alar plate develops into the thalamus and hypothalamus, the roof plate forms the choroid plexus and caudally the epiphysis, with the third ventricle lying within these structures. The telencephalon forms the cerebral hemispheres, developing laterally from the prosencephalon from the 5th week of gestation and contain the lateral ventricles that communicate with the 3rd ventricle through the interventricular foramen, with thickening of the hemisphere wall above to form the hippocampus [1]. The cerebral hemispheres, separated by the longitudinal fissure, develop in anterior, dorsal and posterior directions to form the frontal, temporal and occipital lobes respectively around the central parietal lobe (Table 1.2). Later rapid growth of the cerebral hemispheres results in a highly convoluted surface with ridges (gyri) separated by fissures (sulci).

Brain lobe	Location	Function
Frontal lobe	Anterior to the parietal lobe, separated by the central sulcus. Superoanterior to the temporal lobe, separated by the lateral sulcus.	Higher executive functions e.g. emotional regulation, attention, planning, problem solving, reasoning. Short-term memory tasks. Control of voluntary movement.
Parietal lobe	Superior to the occipital lobe, separated by the parieto-occipital sulcus. Posterior to the frontal lobe, separated by the central sulcus.	Integrating sensory information, including touch, temperature, pain and pressure.
Temporal lobe	Inferioposterior to the frontal lobe, separated by the lateral sulcus.	Processing memories Integrating and processing sensory information, including taste, vision, hearing and language.
Occipital lobe	Posterior to the parietal lobe, separated by the parieto-occipital sulcus. Behind the temporal lobe.	Visual processing centre.
Insula	Beneath the cortex, where frontal, parietal and temporal lobes join.	Processing taste, pain and visceral sensations. Vestibular functions.
Limbic lobe	Medial surface of the cerebral hemispheres surrounding the corpus callosum, including the hippocampus.	Learning and memory. Modulation of visceral, autonomic functions and emotions.

Table 1.2: Lobes of the human cerebrum

Cerebral lobe anatomical boundaries, location within the brain, and function attributed to each lobe [5].

The cerebral hemispheres consist of three distinct regions involving (i) the deep nuclei called the basal ganglia, which play a role in integration of information to fine tune motor control and learning, executive functions, emotions and behaviours [5], (ii) the inner white matter regions which contain afferent and efferent cortical fibres, for example the corpus callosum [3], and (iii) the outer cerebral cortex (grey matter) responsible for consciousness, thought, intellect and memory [5].

1.1.4 Cerebral cortical development

The cerebral neocortex is the largest and most complex part of the brain, with dramatic radial architecture of neurons forming early in development [6]. Three distinct but temporally overlapping processes occur during cortical development [7]:

1. Cell proliferation

Neuroblasts are polarised, consisting of a basal and apical side, and initially divide symmetrically to proliferate within the ventricular zone (the most apical cell layer that lines the ventricles and contains the progenitor cell bodies) [6, 8]. Later neuroblasts form radial glial cells (RGCs), which asymmetrically divide to self-renew and generate a projecting neuron (Figure 1.2) [9, 10]. The basal processes of RGCs progressively elongate to form the radial scaffold framework for neuronal migration [11, 12]. RGCs give rise to neurons and two types of glial cells: astrocytes and oligodendrocytes, with each RGC typically forming just one cell type [13]. They can also produce intermediate progenitors within the subventricular zone (SVZ) which divide symmetrically to produce two differentiated neurons (Figure 1.2) [9, 14, 15].

2. Neuronal migration

Waves of neurons from RGCs in the ventricular zone undergo a complex migration process through the intermediate zone to a subpial position called the cortical plate where they differentiate into mature neurons (Figure 1.2) [16-18], leading to neocortex formation. Migration of pyramidal excitatory cells depends on the RGC process acting as a longitudinal scaffold for locomotive movement of the cell to the cortical plate, called the radial unit hypothesis for cerebral cortical development (Figure 1.2) [6, 19]. Regulation of neuronal migration includes the expression of Reelin protein from Cajal-Retzius cells in the marginal zone and gap junctions between the cells that influence adhesion properties and cytoskeletal dynamics [20-22].
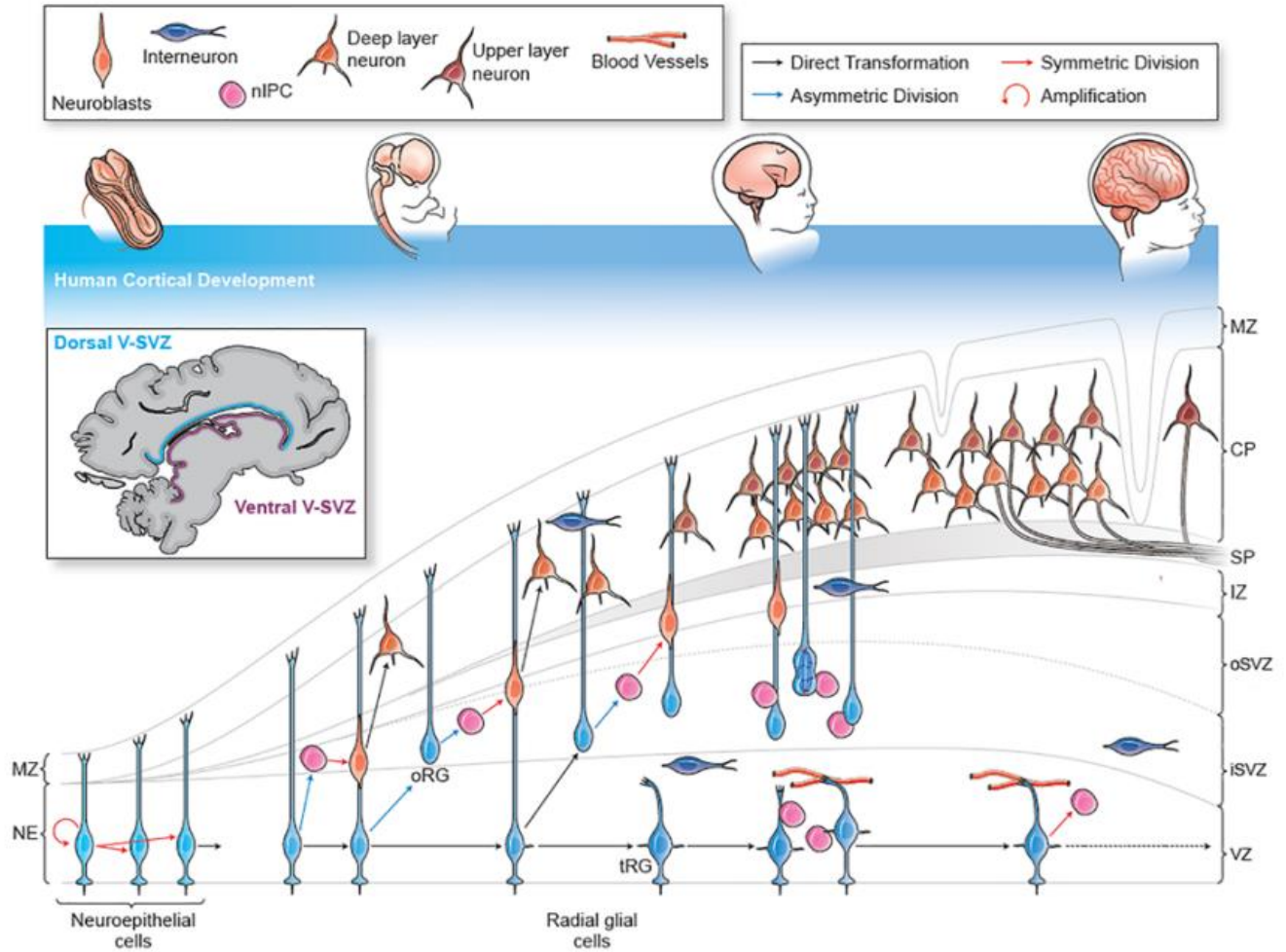
Figure 1.2: Normal development of the human cerebral cortex from early embryo (left) to the neonatal period (right)

In the first trimester neuroepithelial cells symmetrically divide to form a progenitor pool of radial glia cells (RGCs) that can produce neurons through asymmetric division or intermediate precursors. These neurons migrate through the intermediate zone (IZ) along the radial glia scaffold to the marginal zone of the cortical plate where they form a radial organisation in an inside out manner. Cortical neurons then develop synaptic connections and the increased number and stresses results in the development of gyri and sulci. Abbreviations: nIPC, intermediate progenitor cell; MZ, marginal zone; NE, neuroepithelium; CP, cortical plate; IZ, intermediate zone; oSVZ, outer subventricular zone; iSVZ, inner subventricular zone; VZ, ventricular zone; RG, radial glia. Adapted with permission from Subramanian *et al* (2020) [23] (from Frontiers under Creative Commons CC-BY licence, v4).

3. Cortical organisation/connectivity

Neuronal differentiation forms six distinct cortical cell layers which have a stratified appearance (Table 1.3) [3].
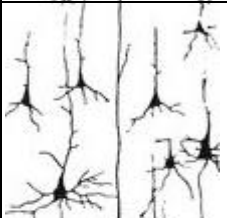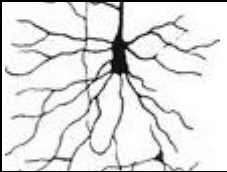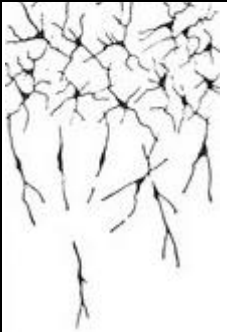
Layers	Golgi stain	Components	Afferents	Efferents
I – Molecular		Axons and dendrites	Intracortical	Intracortical
II – External granular		Densely packed stellate cells and small pyramidal cell	Intracortical	Intracortical
III – External pyramidal		Loosely packed stellate cells and medium pyramidal cells	Intracortical and brainstem	Association and commissural fibres
IV – Internal granular		Densely packed stellate cells only	Intracortical, brainstem and thalamus	
V – Internal pyramidal		Large pyramidal cells, including giant Betz cells	Brainstem	Projection fibres – basal ganglia, thalamus, brain stem, spinal cord
VI - Multiform		Multiple sized pyramidal cells and loosely packed stellate cells		Association and projection fibres - thalamus

Table 1.3: Layers of the human cerebral cortex including components of each layer, afferent and efferent fibres

Golgi staining to show neurons including dendritic projections. Image adapted with permission from ten Donkelaar *et al* (2014) [24] (licence granted by Springer Nature, Appendix 7).

Functional neuronal connectivity is achieved by neuronal differentiation with axonal growth and synapse formation followed by synaptic pruning [25, 26]. Synaptogenesis begins at 22 weeks gestation, although the majority of neuronal

connectivity develops postnatally within the first two years [27]. Neuronal connectivity first follows the radial pattern of cortical neuronal migration from deeper to more superficial layers (Table 1.3) [19] followed by horizontal (tangential) connections. New tangential connections occur in superficial cortical layers I to IV resulting in expansion of these layers and cortical compression that leads to cortical folding beginning at 23 weeks gestation [26]. Primary and secondary sulci form prenatally, with tertiary sulci forming postnatally alongside short association fibres. Synaptic pruning and remodelling occur during cortical development and continue into adolescence, with cortical plasticity in humans occurring throughout life.

Glial cells also form an important part of cortical development and include astrocytes, oligodendrocytes and microglial cells. The number of these cells greatly increases in the third trimester once excitatory neuronal migration is complete [28]. Astrocytes have numerous processes that interact with cells of all types, guide migration of developing neurons and axons and interact with capillaries to ensure neurons have an adequate metabolic and oxygen supply. Oligodendrocytes provide structural support and produce myelin to facilitate axonal conduction with subsequent white matter development. Microglial cells control neuronal proliferation, differentiation and synaptic modification [28].

The process of cortical development from neuroprogenitor cells is regulated by many cellular mechanisms (Table 1.4) involving numerous molecular signalling pathways.

Cellular processes involved in cortical development	Example pathways
Regulation of cell-cycle progression: The transition of neuroblasts to fate-restricted RGCs is associated with an increase in cell cycle length [29].	Ras-MAPK-ERK [30] PI3K-AKT-mTOR [31] P53 signalling [32]
Regulation of cell division machinery: The organisation of microtubules, centrosomes and mitotic spindle formation [33].	Cyclin-CDKs [34]
Regulation of cell differentiation: Differentiation of neuroblasts/RGCs into neurons or glia [23].	JAK-STAT [35]
Apoptosis: Programmed cell death is an important process that controls the size and shape of the developing brain [36].	PI3K-AKT-mTOR [37]
Apical–basal polarity: Controls whether cell division will be symmetrical or asymmetrical and therefore proliferative or neurogenic respectively [9].	TGFβ-SMAD [38] Wnt signalling [39]
Interkinetic nuclear migration: The process of migration of the cell nucleus during the cell cycle in neuroblasts and RGCs [40].	Notch signalling [41]
Primary cilia function: Microtubule-based cilia located on the apical surface of neuroprogenitor cells transduce molecular signalling pathways, assist in maintenance of apical basal polarity and cell cycle regulation [42].	SHH signalling [43] Notch signalling [43]
Membrane trafficking: Golgi-derived membrane trafficking is required for neuronal progenitor proliferation [44].	Phosphoinositide signalling [45]
CSF-derived cues and signalling pathways: The CSF contains diffusible molecules that act as regulators of cortical development [46].	Growth factor signalling [46] Tyrosine kinases [47]
Local cues: Direct cellular contacts influence neurogenesis, including gap junctions and signalling with extracellular matrix components [48].	Reelin signalling [20] Integrin signalling [48] Dystroglycans [49]

Table 1.4: Cellular processes and examples of molecular signalling pathways involved in the regulation of cerebral cortical development

Abbreviations: MAPK; mitogen activated protein kinase, ERK; extracellular signal-regulated kinase, RGC; radial glial cell, PI3K; phosphatidylinositol 3-kinase, AKT; protein kinase B, mTOR; mechanistic target of rapamycin, p53; tumor protein 53, CDK; cyclin dependent kinases, JAK; Janus kinase, STAT; signal transducer and activator of transcription, TGFβ; transforming growth factor beta, SMAD; small mothers against decapentaplegic, Wnt; Wingless/Integrated, SHH; sonic hedgehog, CSF; cerebrospinal fluid.

Specific molecular pathways involved in some of these processes are discussed in detail in sections 1.3.4 and 1.3.5.

1.2 Occipitofrontal circumference

Head circumference or OFC, considered in terms of age and sex, correlates well with intracranial volume (and therefore brain size) and cognitive function, and can be a helpful indicator of underlying brain pathology [50]. The range of OFC considered normal is within greater than -2 standard deviations (SD) and less than +2SD of the mean, although others define the normal range between greater than -3SD and less than +3SD of the mean [51]; within this normal range there is little correlation between OFC and intelligence quotient (IQ). Measurements outside this range are a useful clinical finding and can indicate underlying brain pathology; a greater degree of deviation is strongly correlated with the presence and severity of intellectual disability [51].

OFC charts standardise measurements based on age, sex and gestation and can be converted to standard deviation scores with normal distribution. Several different OFC charts exist based on data from different populations, although population specific charts are not available for some groups and may not be fully representative of a certain population [52]. A growth reference is an anthropometrical summary of statistical measurements from a reference group of children, whereas a growth standard selects children based on their health to represent a standard of how healthy children ought to grow [53]. The UK1990 growth reference OFC chart includes data from children in the UK within the age ranges of 23 weeks gestation to 18 years in male and 17 years in females [54]. In 2007 the World Health Organisation (WHO) produced international growth standard OFC charts from birth to five years, which include children from the United States, Brazil, Ghana, India, Oman and Norway [55]. The UK1990 and WHO charts were combined in 2009 to create the UK-WHO chart, based on the growth of breastfed infants from the ages of 32 weeks gestation to 18 years. OFC charts are primarily available up to four years, and while charts after 4 years are available these do not correlate with the 0-4 year charts with a shift in the reference data at this age.

Studies comparing the WHO standard OFC charts to measurements from 55 different populations found that measurements were consistently below the mean (0.5-1 SD) [56]. Due to this huge variation, it has been suggested that international 'standard' OFC charts should not be used across all populations as there is a considerable risk of an incorrect diagnosis of microcephaly or macrocephaly, depending on an individuals' ethnic origin. Therefore many countries have developed population specific growth charts and, where possible, the most ethnically suitable growth chart should be used to record OFC with the same chart being used for all measurements for a child [51]. Additionally head circumference and brain size in children and adults is known to be correlated with height [57]. British OFC charts for adults based on the height of an individual have been produced from a relatively small number of healthy individuals [58]. It can be misleading to consider an OFC measurement for an individual without knowing their height. Further, as most paediatric OFC charts of stop at 16-18 years errors may occur, particularly in males who may often undergo a later pubertal growth spurt than females [58]. An abnormality in head circumference can be described as proportionate or non-proportionate based on the context of the affected individual's body stature. Proportionate means head size is normal in relation to stature. It is also important to measure parental OFC, particularly in children where this is an isolated abnormality, as up to 50% of normal OFC variation is familial [59].

1.3 Abnormalities of brain growth and development

Disruption of any of the neurodevelopmental processes of neuroprogenitor cell proliferation, differentiation, migration, organisation and connectivity, may result in abnormalities of brain growth and development, including altered brain size, structure and aberrant neuronal organisation [23]. Causes include a wide range of environmental (infective or teratogenic agents and acquired brain injuries) and genetic (germline or somatic single gene variants, epigenetic, metabolic, mitochondrial, repeat expansion diseases or chromosomal aberrations including copy number variants (CNVs)) causes.

1.3.1 Monogenic neurodevelopmental disorders

These are a large group of clinically and genetically heterogeneous conditions affecting the development, structure and functioning of the central nervous system (CNS). Whilst individually rare, inherited neurodevelopmental disorders as a group significantly contribute to the burden of neurodevelopmental impairment. Determining an underlying genetic diagnosis for inherited neurodevelopmental disorders is challenging due to this degree of genetic and phenotypic heterogeneity. Many inherited neurodevelopmental disorders are highly pleiotropic with multiple seemingly unrelated features affecting different organ systems, for example ciliopathy disorders like Joubert syndrome characterised by ophthalmic, renal, neurological, skeletal and behavioural features [60]. Variants in different genes can cause similar phenotypic features, for example Fanconi anaemia is associated with biallelic variants in *BRCA1*, *BRCA2*, *BRIP1*, *ERCC4*, *FAAP100*, *FANCA*, *FANCB*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, *FANCI*, *FANCL*, *FANCM*, *PALB2*, *RAD51*, *RAD51C*, *RFWD3*, *SLX4*, *UBE2T*, *XRCC2* and *XRCC9* [61]. Different variants within the same gene can cause several different phenotypes for example variants in *Filamin A* (*FLNA)* can result in distinct neuronal migration, skeletal or cardiac disorders [62] making it challenging to identify the precise molecular diagnosis. Due to this clinical and genetic heterogeneity, genetic studies of neurodevelopmental disorders need to be unbiased and genome-wide, including utilising testing technologies such as genomic microarray, next generation sequencing (NGS) and targeted gene panels, gene agnostic hypothesis free whole exome or whole genome sequencing (WES/WGS) [63]. Recent advances in genomic NGS technologies over the last ~10 years have led to an explosion in novel monogenic rare disease gene discovery [64] (2841 genes are listed in the sysNDD database (https://sysndd.dbmr.unibe.ch/), accessed June 2022), although up to 50% of single gene disorders are thought to remain to be undiscovered [65, 66]. Monogenic neurodevelopmental disorders can be associated with microcephaly, macrocephaly or normocephaly.

1.3.2 Microcephaly

The term microcephaly means "small head size" and is defined as an occipitofrontal circumference (OFC) less than or equal to two standard deviations

below the mean for an individual's sex and age (although most define 'severe' or clinically significant microcephaly as an OFC greater than or equal to three standard deviations below the mean) [51]. The term microencephaly relates directly to the size of the patient's brain. When microcephaly is present at birth (prenatal in onset) this is termed primary microcephaly and is usually the result of impaired brain growth during embryonic and/or fetal development. Secondary microcephaly develops after birth (postnatal in onset) and is usually the result of a degenerative process. Progressive microcephaly refers to the severity of microcephaly increasing over time, so OFC measurements cross the centile lines and become progressively smaller, this is more commonly seen with secondary microcephaly. Causes of microcephaly can be divided into acquired (for example intrauterine infection, teratogens, stroke) or genetic [67]. Microcephaly is a common phenotypic feature that can be an isolated finding (non-syndromic) or associated with other phenotypic features (syndromic) where it can be a core characteristic of a condition or a secondary non-mandatory feature. Microcephaly is associated with 1778 clinical syndromes in Online Mendelian Inheritance in Man (OMIM) (https://www.omim.org/) [68] with huge genotypic and phenotypic heterogeneity.

Worldwide the incidence of primary microcephaly is 1.3 to 150 per 100,000 live births [69]. There are a vast number of genetic causes of primary microcephaly with a wide range of underlying pathogenic processes; Table 1.5 provides a non-exhaustive list.

Type of inherited disorder	Examples
Chromosomal disorders	4p16.3 deletion (Wolf-Hirschhorn syndrome) [70] 7q11.23 deletion (Williams syndrome) [71] Microdeletion/duplication of specific microcephaly-related genes
Non-syndromic autosomal recessive primary microcephaly (MCPH)	Centrosomal proteins e.g. *ASPM* [72], *WDR62* [73], *CENPJ* [74] Cytokinesis proteins e.g. *CIT* [75]*, KIF14* [76] Apoptotic processes e.g. *WDFY3* [77]*, ZNF335* [78]
Microcephaly with dwarfism	Microcephalic osteodysplastic primordial dwarfism (MOPD) I and II [79, 80] Seckel syndrome [81] Meier-Gorlin syndrome [82]
DNA repair disorders	Nijmegen breakage syndrome [83] Bloom syndrome [84] *PNKP*-related microcephaly [85]
DNA replication disorders	Meier-Gorlin syndrome [82]
Ciliopathies	Centrosomal proteins e.g. *ASPM* [72], *WDR62* [73]*, CENPF* [86] Meckel syndrome [87] Microcephaly and chorioretinopathy e.g. *KIF11* [88]*, PLK4* [89]
Metabolic disorders	Serine deficiency disorders [90] Amish lethal microcephaly (alpha ketoglutaric aciduria) [91]
Syndromic single gene disorders	Holoprosencephaly (HPE1-13) [68] Cornelia de Lange syndrome [92]

Table 1.5: Examples of genetic causes of primary microcephaly

Secondary microcephaly occurs postnatally, may be syndromic or non-syndromic, and affected individuals typically display progressive motor and cognitive deterioration with or without seizures. There are a large number of genetic conditions associated with secondary microcephaly, which can be grouped based on pathogenesis (Table 1.6).

Type of inherited disorder	Examples
Chromosomal disorders	17p13.3 deletion (Miller-Dieker syndrome) [93]
Metabolic disorders	Congenital disorders of glycosylation [94] Mitochondrial disorders [95] Peroxisomal disorders [96] Amino acidopathies/organic acidurias [97, 98] Glucose transporter defect [99] Menkes disease [100]
Syndromic single gene disorders	Aicardi-Goutieres syndrome [101] Cockayne syndrome [102] Rett syndrome [103]
Membrane trafficking disorders	Cohen syndrome [104] TRAPPopathies [105] Golgipathies [106] Periventricular heterotopia with microcephaly (*ARFGEF2*) [44]
Neuronal migration/connectivity disorders	Tubulinopathies [107] Lissencephalies e.g. DCX [108], *TUBA1A* [109], *NDE1* [110] Polymicrogyria e.g. *TUBB2B* [111]
Amino acid or protein synthesis	tRNA synthetase deficiencies *QARS* [112], *AARS* [113], *CARS1* [114]

Table 1.6: Examples of genetic causes of secondary microcephaly

1.3.2.1 Autosomal recessive microcephalic disorders

Neurodevelopmental disorders in individuals from endogamous families (including Amish and Pakistani communities described herein) are most frequently caused by autosomal recessive inheritance of pathogenic founder genetic variants (discussed in detail in section 1.6) [66]. Therefore, this thesis focuses on autosomal recessive monogenic disorders of brain growth and development.

Table 1.4 summarises the cellular processes and signalling pathways involved in cortical neurodevelopment. Disruption of several of these processes/pathways can affect neurogenesis and the proliferation of RGCs within the developing brain, resulting in fewer mature neurons and a reduced brain size [115, 116] and have been associated with autosomal recessive microcephalies, with several examples described below. Following a literature review of genes listed in OMIM, sysNDD (https://sysndd.dbmr.unibe.ch/), PanelApp (https://panelapp.genomicsengland.co.uk/) and Pubmed

(https://pubmed.ncbi.nlm.nih.gov/), 394 genes were identified to be associated with autosomal recessive microcephaly. These are summarised in Figure 1.3, grouped into conditions affecting the same cellular process or localisation within the cell. This clearly demonstrates the huge genetic heterogeneity of autosomal recessive microcephaly disorders.

A specific group of non-syndromic primary microcephaly disorders have been defined as MCPH (microcephaly primary hereditary) and typically show autosomal recessive inheritance (grey circle in Figure 1.3). Individuals with MCPH usually have some degree of intellectual disability with speech delay and variable motor delay [117], although cognitive function is often relatively preserved despite a marked reduction in cortical volume [118]. To date, 25 genes have been identified as autosomal recessive MCPH genes (Figure 1.3, grey circle) [68, 119] of which 21 encode proteins involved in cell cycle regulation, cell division, centrosome biogenesis or apoptotic processes (Figure 1.3 overlapping red circle). Many of the centrosomal MCPH proteins interact with each other in centriole biogenesis to enable centriole duplication, a crucial process in neurogenesis [116]. Other MCPH proteins control processes of DNA repair, chromosomal stability, membrane trafficking, cytokinesis and primary ciliary function. For example, *ASPM (abnormal spindle-like, microcephaly-associated)*, the commonest genetic cause of primary microcephaly [120, 121], plays a role in centriole biogenesis, the process of cytokinesis, and primary ciliary function [116] and mediates some of these functions through positive regulation of the Wnt signalling pathway [122].

Receptors/cell surface anchor/ECM

Amino acid/protein synthesis

	AARS1	VARS1	
NAA20	CARS1	YARS1	ASNS
CTU2	FARSA	ADAR	AIMP1
EPRS1	FARSB	ADARB1	AIMP2
WDR4	FARS2	ADAT3	ALDH18A1
OSGEP	IARS	CLP1	TRMT1
TP53RK	KARS1	DHFR	TRMT10A
TPRKB	NARS1	NSUN2	SEPSECS
PSAT1	QARS	GPT2	PYCR1
	SARS1	PRMT7	PYCR2

Congenital disorders of glycosylation

COG1	COG8	DPM1	TMEM165
COG2	ALG1	NGLY1	SLC35A1
COG4	ALG3	NUS1	SLC35A3
COG5	ALG9	GALNT2	SLC35C1
COG6	ALG12	STT3A	RFT1
COG7	DPAGT1	FUT8	

Dystroglycanopathy

POMT1	POMGNT1	FKRP
POMT2	GMPPB	

Other

SLC25A19	HTRA2	PLPBP	SC5D
PHGDH	CYB5R3	ALDH6A1	COASY
FTO	DHCR7	FAR1	PAH
AMPD2	CLPB	PDHX	

Metabolic

Mitochondria

SLC25A22	TRIT1	HPDL
FBXL4	MTFMT	SDHD
MICU1	MRPS34	C2orf69
SFXN4	HSP60	SHMT2
MFF	TMEM70	MICOS13
NDUFV2	PMPCB	ELAC2
RARS2	GFM2	TMX2
VARS2	PISD	RNASET2
	GOT2	

NALCN	SCN1B	OCLN
UNC80	CNNM2	PCDH12
SLC1A4	KCNMA1	PCDHGC4
SLC12A5	CACNA1B	NTNG2
SLC12A2	IGF1R	PIGY
SLC38A3	IGF1	ITGAV
MFSD2A	GTPBP2	ENTPD1
LINGO1	GRM7	PI4KA
ST3GAL5	ATP1A2	

Nucleus/transcription factors

	PPP1R15B	
	METTL5	PUS3
MED17	RNPC3	PUS7
MED23	ZNF526	PRDM15
MED25	ZNF407	TASP1
MED27	ZNF699	NUP188
TAF2	GTF2E2	INTS8
TAF6	ZBTB11	LARP7
TAF13	P4HTM	RNU7-1
DEAF1	ACTL6B	BRF1
	MAB21L1	ZPR1
	ZNHIT3	

ER/protein degradation

UBE3B	PLAA
UFM1	KLHL7
UFC1	CCDC47
UFSP2	SMG8
UBR1	SMPD4
UBR	MBOAT7
EFL1	MSMO1
PIGH	EMC1
TOR1A	DCPS
EMG1	TMTC3
	AGMO

MCPH

RRP7A

CENPE
CDK6

ASPM	ZNF335	KIF14
WDR62	ANKLE2	CIT
CDK5RAP2	KNL1	NUP37
STIL	NCAPD3	
CEP135	NCAPD2	
SASS6	NCAPH	

Membrane trafficking/intracellular transport

COPB1	TRAPPC2L	PCLO	VPS11	
SCYL2	TRAPPC4	AP3B1	VPS13B	PGAP2
ABCA2	TRAPPC6B	AP3B2	VPS50	STAMBP
DYM	TRAPPC9	AP4B1	VPS51	EXOSC3
IQSEC1	TRAPPC10	AP4E1	VPS53	HIKESHI
EXOC7	TRAPPC11	AP4M1		IPORIN
	TRAPPC12	AP4S1	TBC1D23	ATP9A
COPB2	COH1	DENND5A	GOLGA2	THOC6
MFSD2A	RAB3GAP1	YIPF5		NAPB
TRAPPC14	RAB3GAP2	SNAP29		DMXL1
	RUSC2	RAB18	NUP214	HID1
	FIG4	TBC1D20	WDR45B	YIF1B

STAMBP
CHMP1A

Cell cycle/division/centrosomal/apoptosis

			CDK10		
		TSEN15	TTC5		
CENPF	NDE1	TSEN54	NCAPG2		
TSEN2	BUB1B	CKAP2L	CAPN15		POC1A
INPP5K	TBCD	WWOX	RBL2	KATNB1	PLK4
INPP4A	TBCE	PTRH2	MPLKIP	TUBGCP6	CEP55
MINPP1	WDR73	NUP107	VRK1	NEURL4	CEP57
	TECPR2	EXT2	MCM7	DYNC1I2	PRUNE1
	SPATA5	TOE1	HHAT		
	OTUD6B		SMO		

MCPH1
PHC1

ESCO2

MOPD

PCNT
RNU4ATAC

Seckel syndrome

NIN CENPJ
CEP63 CEP152

Meier-Gorlin syndrome

ATR
TRAIP
NSMCE2
DNA2
RBBP8
ATRIP

ORC1
ORC4
ORC6
CDT1
CDC6
CDC45L
MCM5

DONSON
PRIM1

Microcephalic primordial dwarfism

DNA repair, replication and chromosomal stability

PNKP	ERCC1	FANCA	BLM
NBN	ERCC2	FANCC	TOP3A
NHEJ1	ERCC3	FANCD2	RMI1
XRCC2	ERCC4	FANCE	RAD50
XRCC4	ERCC5	FANCF	RAD51C
XRCC9	ERCC6	FANCG	BRCA1
PCNA	ERCC6L2	FANCI	BRCA2
DDX11	ERCC8	FANCL	BRIP1
RNASEH2A	LIG4	SLX4	PALB2
RNASEH2B	TREX1	MAD2L2	BRAT1
RNASEH2C	TELO2	RFWD3	
SAMHD1	XPA	RTEL1	
ZNF668	PARN	TDP2	
TRIP13	PRKDC		
MRE11A	TTI2		

Cytoplasm

NCDN
LINS1
ITPA
UGP2
UBA5
SPATA5L1
AGTPBP1

Cytoskeleton/microtubule

POC1A
KATNB1 PLK4
KIFBP
DIAPH1
CCDC88A
CRIPT
BCAS3
CTNNA2
ADD3
PLEKHG2
TUBGCP2
TUBGCP4
FRMD4A
SVBP
RTTN
PTPN23
WDR11
CCDC32

Figure 1.3: Schematic diagram of 394 genes associated with autosomal recessive microcephalic disorders identified by literature review

Autosomal recessive microcephaly genes were identified following literature review of Pubmed, OMIM, sysNDD and PanelApp databases (accessed 29/04/2022). Genes are grouped based on function or cellular localisation and demonstrates the huge genetic and phenotypic heterogeneity of disorders with features of microcephaly. Green text indicates the presence of 69 high confidence disease genes ("green genes") in Genomics England PanelApp [123]. Genes in italic indicate association with primary ciliary function, termed ciliopathies. Abbreviations: MCPH, microcephaly primary hereditary; ER, endoplasmic reticulum; ECM, extracellular matrix; MOPD, microcephalic osteodysplastic primordial dwarfism.

A number of other genes associated with autosomal recessive microcephaly encode proteins involved in cell cycle progression/cell division, centrosome biogenesis and apoptotic processes (Figure 1.3, red circle), DNA repair, replication and chromosome stability (purple circle) which overlap with those causing microcephalic primordial dwarfism (black dashed line), membrane trafficking/intracellular transport (green circle), microtubule/cytoskeleton structure (blue circle) and amino acid or protein synthesis (orange circle). Many of the functional processes in these groups overlap, for example centriole function and microtubule structure, thus many genes lie within more than one group (Figure 1.3). Autosomal recessive metabolic disorders (inborn errors of metabolism) form a large group of 41 genes associated with microcephaly (Figure 1.3, red shaded square). These can be categorised into several subgroups, the largest of which is congenital disorders of glycosylation, dystroglycanopathies, peroxisomal disorders and organic acidurias. Mitochondrial disorders due to biallelic variants in nuclear encoded genes are also associated with microcephaly (brown shaded circle).

1.3.3 Macrocephaly/megalencephaly

The term macrocephaly means "large head size" and is defined as an OFC more than or equal to two standard deviations above the mean for an individual's sex and age. Although most experts define 'severe' or clinically relevant macrocephaly as an OFC more than or equal to three standard deviations above the mean, as many individuals with mildly enlarged head size (2-3 SD) have normal development [124]. Megalencephaly refers specifically to a large brain size of greater than 2 to 3 SD above the mean. It is important to distinguish macrocephaly and megalencephaly, as macrocephaly can have other causes in addition to a large brain size, including environmental factors (infection, teratogens), skull bone structure abnormalities, intracranial masses, subdural fluid collection, hydrocephalus and arteriovenous malformations. Whereas, megalencephaly is almost exclusively caused by genetic factors [125].

Onset can occur early in fetal development, termed prenatal (congenital) megalencephaly, or can develop more slowly during infancy and early childhood, termed postnatal megalencephaly. However, it is more common for

megalencephaly to be classified into two subtypes (i) metabolic and (ii) anatomic based on the underlying pathogenesis [124]. Megalencephaly may occur as part of a clinical syndrome (syndromic) or an isolated feature (non-syndromic) sometimes with other associated neurological features. Benign familial megalencephaly refers to an abnormally large brain with no neurological impairment, which is typically familial [126]. Megalencephaly and/or macrocephaly display notable genotypic and phenotypic heterogeneity and are associated with 574 clinical syndromes in OMIM [68].

Metabolic disorders associated with megalencephaly are characterised by several different pathophysiological mechanisms including disruption of enzyme function, biochemical processes, and abnormal accumulation of substrates in the brain often affecting a specific cell type and include a wide range of disorders that can be divided into three broad groups [126] (Table 1.7).

Metabolic disorder	Examples
Defects of organic acids	Glutaric aciduria type 1 (*GCDH*) [127] L2 Hydroxyglutaric aciduria (*L2HGDH*) [128] D2 Hydroxyglutaric aciduria (*D2HGDH*) [129]
Leukoencephalopathies [130]	Canavan disease (*ASPA*) [131] Alexander disease (*GFAP*) [132] Leukoencephalopathy with vanishing white matter (*EIF2B1-5*) [133] Megalencephalic leukoencephalopathy with subcortical cysts e.g. *MLC1, HEPACAM* [134]
Lysosomal storage diseases	Tay-Sachs disease (*HEXA*) [135] Sandhoff disease (*HEXB*) [136] Mucopolysaccharidosis e.g. Hunter (*IDS*) [137], Hurler (*IDUA*) [138], Sanfilippo syndromes (*SGSH, HAGLU, ARSB*) [139, 140] Krabbe disease (*GALC*) [141]

Table 1.7: Inherited metabolic disorders associated with megalencephaly

Anatomic (developmental) megalencephaly includes a vast number of genetic conditions that can be grouped based on the underlying physiological process or signalling pathways disrupted (Table 1.8). Anatomic megalencephaly can be caused by germline genetic conditions associated with generalised overgrowth of both the brain and body, for example Sotos, Simpson-Golabi-Behmel and

Cowden syndromes, or isolated brain overgrowth, for example disorders of the PI3K-AKT-mTOR pathway (mTORopathies), Noonan syndrome and neurofibromatosis (NF1) [126] (see Table 1.8 for specific gene examples). Several megalencephalic disorders are caused by somatic mutation in a 'second hit' hypothesis, which occurs when a *de novo* mutation arises after zygote formation, so only affecting the cells that arise from the mutated cell in a specific region/proportion of an individual depending on the developmental stage this occurs. This can lead to a clinical presentation where there is overgrowth of only a proportion of the brain, for example hemimegalencephaly as seen in *mTOR*, *PIK3CA* and *AKT3*-related conditions [142], or the cortical tubers observed in tuberous sclerosis [143]. Individuals with hemimegalencephaly usually display severe neurological impairment with early onset seizures refractory to medication [144].

Type of inherited disorder	Examples
Chromosomal disorders	Microdeletion/duplication of specific megalencephaly-related genes e.g. *NSD1* deletion [145]
Disorders of signalling pathways:	
• mTORopathies [146]	Cowden/*PTEN*-hamartoma tumor syndrome [147] *PIK3CA*-related overgrowth disorders [142] Megalencephaly–polymicrogyria–polydactyly–hydrocephalus (MPPH) syndrome e.g. *AKT3 , CCND2, PIK3R2* [142] Smith-Kingsmore syndrome (*mTOR*) [148] Tuberous sclerosis (*TSC1/2*) [31]
• Rasopathies	Noonan syndrome e.g. *PTPN11, NRAS, RAF, SOS1* [149] Neurofibromatosis type 1 (*NF1*) [150] Cardiofaciocutaneous syndrome e.g. *BRAF, KRAS* [151] Costello syndrome (*HRAS*) [152]
• SHH pathway [153]	Greig cephalopolysyndactyly syndrome (*GLI3*) [154] Acrocallosal syndrome (*KIF7*) [155] Gorlin syndrome e.g. *PTCH1, PTCH2* [156, 157]
• Notch signalling	Simpson-Golabi-Behmel syndrome type 1 (*GPC3*) [158] 1q21.1 duplication syndrome [159]
Epigenetic/transcription regulators	Sotos syndrome (*NSD1*) [145] Weaver syndrome (*EZH2*) [160] Tatton-Brown-Rahman syndrome (*DNMT3A*) [161] Malan syndrome (*NFIX*) [162]
Ciliopathies	Simpson-Golabi-Behmel syndrome type 2 (*OFD1*) [163] Greig cephalopolysyndactyly syndrome (*GLI3*) [154]
Receptor tyrosine kinase disorders	Robinow syndrome (*ROR2*) [164] Achondroplasia (relative macrocephaly) (*FGFR3*) [165]

Table 1.8: Genetic causes of anatomic megalencephaly

1.3.3.1 Autosomal recessive megalencephalic disorders

While inherited microcephalies display predominantly autosomal recessive inheritance, most developmental megalencephalic disorders show autosomal dominant inheritance [126]. Therefore, these genes likely play such a crucial role in brain development that complete absence or loss of function would be incompatible with life. Disruption of several signalling pathways involved in cortical neurodevelopment (Table 1.4) are known to result in abnormalities of neurogenesis with an increase in the number or size of cells, resulting in megalencephaly. Figure 1.4 summarises a total of 78 genes associated with

autosomal recessive megalencephalic disorders, identified following a literature review of genes listed in OMIM, sysNDD, Pubmed and PanelApp databases, grouped according to function or subcellular localisation.

Genetic variants within the PI3K-AKT-mTOR pathway entail the commonest cause of megalencephaly [166] (Figure 1.4, purple box) for example polyhydramnios, megalencephaly and symptomatic epilepsy (PMSE) caused by biallelic germline variants in *STRADA* [167]. This pathway, discussed in detail in Chapter 4, plays a critical role in several processes involved in brain development, including neurogenesis, cell cycle regulation, cell differentiation, apoptosis, neuronal migration and synaptogenesis [37, 146]. Further groups of autosomal recessive megalencephaly genes include those involved in cell cycle, cell division, stability and apoptosis (Figure 1.4, blue circle) and ciliopathies (yellow circle) that encompass several genes involved in SHH signalling (green box) and overlap with genes involved in membrane trafficking disorders (red box) and cytoskeletal regulation (blue circle). Several genes play a role in multiple processes, for example *KIF7* has been identified to play a role in SHH signalling, microtubule dynamics and primary cilia function [155] and is associated with Acrocallosal syndrome/Joubert syndrome 12, a ciliopathy disorder characterised by severe intellectual disability, neonatal hypotonia, megalencephaly, dysmorphic features and polydactyly [155].

The metabolic megalencephalies display mostly autosomal recessive inheritance and can be divided into three main types: organic acidurias, lysosomal storage disorders and leukoencephalopathies, although several do not fit into these groups (Table 1.7, Figure 1.4, red shaded square) [126]. Two genes have been associated with recessive mitochondrial disorders (brown shaded circle) [168].

Finally, it is interesting to note that some genes have been associated with both microcephaly and megalencephaly/hemimegalencephaly (for example, *AKT3 NOTCH2NL*, *MYCN*, *NSD1* and 10q22-23, 16p11.2 loci [125, 142, 169]*), depending on whether a variant is activating or loss of function.
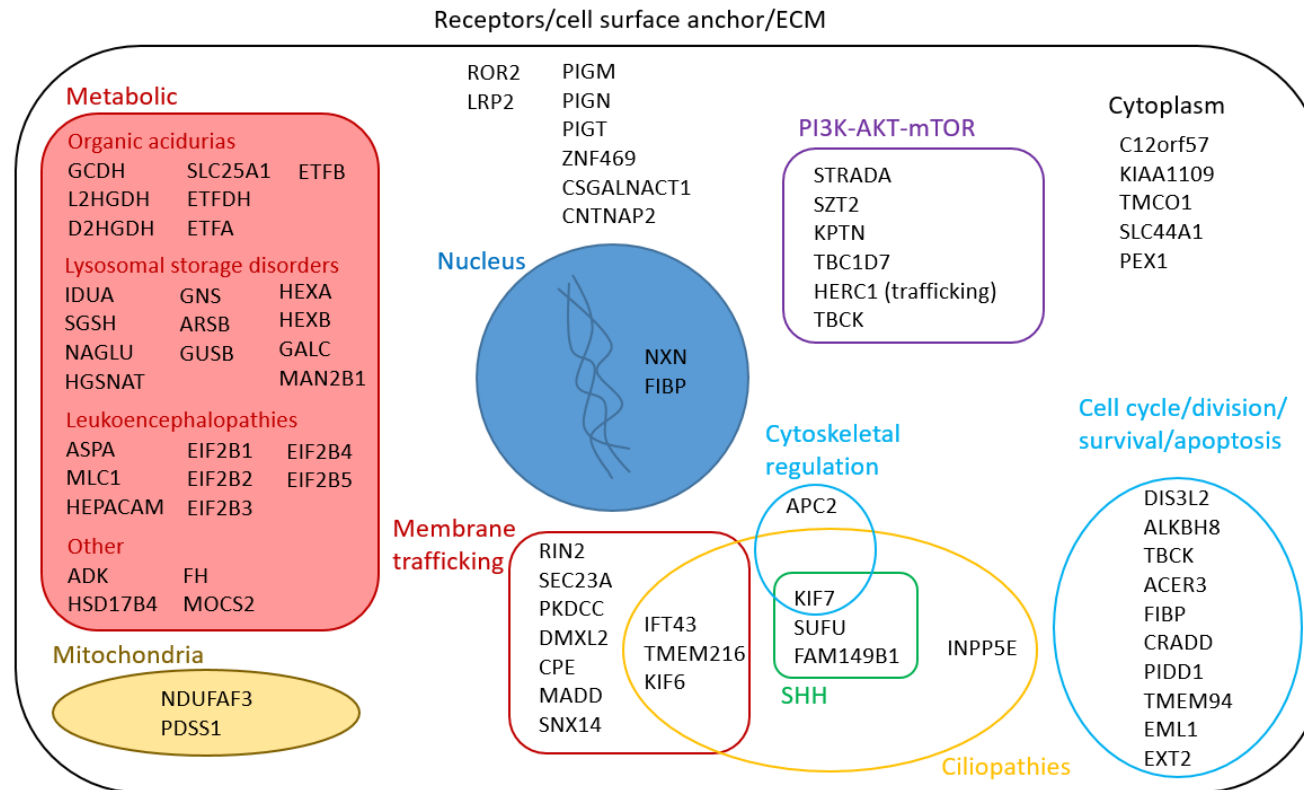
Figure 1.4: Schematic diagram of 78 genes associated with autosomal recessive megalencephalic disorders identified by literature review

Autosomal recessive megalencephaly genes identified following literature review of Pubmed, OMIM, sysNDD and PanelApp databases (accessed 29/04/2022). Genes are grouped based on function or cellular localisation and demonstrates the huge genetic and phenotypic heterogeneity of disorders with features of megalencephaly. Disorders associated with skeletal dysplasia/abnormalities with relative macrocephaly, often with no or few neurodevelopmental phenotypes, have not been included in this summary as this macrocephaly likely represents part of a skeletal defect or hydrocephalus rather than true megalencephaly. Abbreviations: ECM, extracellular matrix; SHH, sonic hedgehog.

1.3.4 Disorders of cortical development

Errors in the processes involved in cerebral cortical development, including neuronal migration or connectivity, result in a large heterogeneous group of cortical malformations with a wide spectrum of phenotypes [170].

Neuronal migration disorders are caused by genetic defects affecting proteins involved in migrational processes. This includes lissencephaly, meaning "smooth brain", a spectrum of conditions characterised by the absence of or abnormality of the sulci or gyri of the brain surface with four-layered cortical structure, in which phenotypical severity correlates with the severity of agyria [171]. Heterotopia, the abnormal clustering of cortical neurons, typically occurs as periventricular or subcortical heterotopia arising from defects impacting the ventricular zone and cortical layers respectively [170]. The "tubulinopathies" entail a further large group of neuronal migration disorders due to mutation of genes encoding tubulin subunits, which may be associated with a range of malformations of cortical development [172]. This is typified by *TUBA1A* mutation, which is associated with microcephaly, variable lissencephaly from agyria to subcortical band heterotopia, polymicrogyria, pontocerebellar hypoplasia and agenesis of the corpus callosum [173]. The focal cortical dysplasias include a spectrum of focal developmental malformations of the cortical cytoarchitecture associated with intractable epilepsy [174] that arise from a 'second hit' *de novo* somatic mutation.

Malformations of cortical connectivity include polymicrogyria, a highly heterogeneous spectrum of malformations including cobblestone dysplasia and schizencephaly, characterised by overfolding of the cerebral cortex with abnormal cortical layering [175]. Polymicrogyria is thought to represent a post-migrational disorder affecting the connectivity of cortical neurons after they have migrated from the periventricular zone [176]. The commissural structures are white matter tracts connecting the two cerebral hemispheres, the largest being the corpus callosum. Disruption of several developmental brain processes including cilia formation, axon guidance, cell adhesion, proliferation and migration have been associated with abnormalities in corpus callosum development, from hypoplasia to complete/partial agenesis [177, 178].

Examples of genetic causes of each of these disorders are summarised in Table 1.9.

Disorders of cortical malformation	Examples
Lissencephalies: an absence or abnormality of the sulci or gyri of the brain	*DCX*, *LIS1* [179], *RELN* [180] 17p13.3 deletion (Miller-Diecker syndrome) [93]
Heterotopias: an abnormal clustering of cortical neurons e.g. periventricular or subcortical	*FLNA* [181], *KIF2A* [182], *MAP1B* [183], *ARFGEF2* [44] and *FAT4* [184]
Focal cortical dysplasias: a spectrum of focal developmental cortical malformations arising from 'second hit' somatic variants	*TSC1/2* [143], *CNTNAP2* [185], *DEPDC5*, *NPRL2* and *NPRL3* [186]
Polymicrogyria: an overfolding of the cerebral cortex with abnormal cortical layering	*TUBA1A* [173], *WDR62* [187], *FIG4* [188] and *NDE1* [110] 1p36.3 and 22q11.2 microdeletions [175]
Commissural/callosal defects: abnormal formation of white matter tracts connecting the cerebral hemispheres e.g. agenesis/hypoplasia of the corpus callosum	Coffin-Siris, Baraitser-Winter and Aicardi syndromes [177] 1q43-q44 microdeletion

Table 1.9: Disorders of cortical malformation

1.3.5 Cerebellar malformations

Cerebellar malformations may occur in isolation, alongside brainstem abnormalities or as part of a syndrome affecting the development of other brain structures [189]. Clinically, cerebellar features include cerebellar ataxia, intention tremor, nystagmus, dysarthria, hypotonia, titubation, vertigo and cognitive impairment. Cerebellar hypo- and hyperplasia refer to either under or overdevelopment of the cerebellum and are non-specific features of many disorders, usually associated with cognitive and motor impairment [189], for example biallelic pathogenic *RELN* variants associated with lissencephaly and severe abnormalities of the cerebellum, hippocampus, and brainstem [180, 190]. Cerebellar dysplasia is defined by abnormal cerebellar foliation, white matter arborisation and grey matter junction, which can affect small focal regions to the whole cerebellum [191], and is often associated with hypoplasia. Pontocerebellar hypoplasia refers to hypoplasia of the cerebellum and brainstem structures, which usually occurs at an earlier stage of brain development [192].

1.4 Neurodevelopmental disorders

Neurodevelopmental disorders are a large group of conditions affecting brain development, structure and function. There are two widely used classification systems: The International Statistical Classification of Diseases and Related Health Problems, 11th edition (ICD-11) [193] produced by the World Health Organisation (WHO) and the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-V) [194].

ICD-11 defines neurodevelopmental disorders as "*behavioural and cognitive disorders that arise during the developmental period that involve significant difficulties in the acquisition and execution of specific intellectual, motor, language, or social functions*" [193] and provides diagnostic groups for neurodevelopmental disorder including:

- Disorders of intellectual development

- Developmental speech and language disorders

- Autism spectrum disorder

- Developmental learning disorder

- Developmental motor coordination disorder

- Attention deficit hyperactivity disorder

- Stereotyped movement disorder

DSM-V classifies neurodevelopmental disorders into groups similar to ICD-11 [194]. Wider conceptualisations of neurodevelopmental disorders include other conditions outwith the behavioural descriptions included in ICD-11 or DSM-V groups but are caused by abnormalities in brain development that result in neurological or psychiatric disorders or syndromes, for example autistic features in Rett syndrome. The ICD-11 classification provides an additional classification for these disorders as 'Secondary neurodevelopmental syndrome: Secondary mental or behavioural syndromes associated with disorders or diseases classified elsewhere' [193]. Although the features within each of the ICD-11 classification groups can occur independently they often show significant phenotypic overlap and comorbidity, and it has been hypothesised that these conditions lie on a continuum of neurodevelopmental

disruption [195]. Wider concepts include any condition caused by disruption of neurodevelopmental processes, for example neuronal migration disorders or epilepsy syndromes, and some have suggested that new classification of neurodevelopmental disorders based on molecular basis/genotype would be more accurate and informative [195]. This thesis uses the term neurodevelopmental disorder to widely refer to any condition that arises from disruption of brain development or affects a child's neurodevelopmental skills.

1.4.1 Intellectual disability

Intellectual disability (previously described as 'mental retardation') is defined by the WHO as "*a significantly reduced ability to understand new or complex information and to learn and apply new skills (impaired intelligence). This results in a reduced ability to cope independently (impaired social functioning), and begins before adulthood, with a lasting effect on development.*"[196].

ICD-11 classifies the severity of intellectual disability as mild, moderate, severe or profound. This is based on an individual's adaptive functioning as well as standardised test scores (Table 1.10). For severe to profound intellectual disability standardised tests of intelligence are not reliable or validated, so classification is based solely on an individual's adaptive behaviour. For children under the age of five a disorder of intellectual development cannot be confirmed, as intellectual functioning or adaptive behaviour is difficult to assess and not reliable and the term global developmental delay should be used instead (see below). There are several different types of neuropsychiatric tests of cognitive function based on the age of an individual, including the Stanford-Binet Intelligence Scale [197], the Wechsler Intelligence Scale for Children 4th Edition (WISC-IV) [198] (6-16 years) or the Wechsler Adult Intelligence Scale 4th Edition (WAIS-IV) [199] (over 16 years), which assess cognitive performance based on a standardised normal distribution in four domains including verbal comprehension (VCI), perceptual reasoning (PRI), processing speed (PSI) and working memory (WMI) indices, which can be combined to generate a full-scale intelligence quotient (FSIQ) score.

Severity	Mild	Moderate	Severe	Profound
Centile	0.1-2.3	0.003-0.1	<0.003	<0.003
Standardisation score	2-3 SD below the mean	3-4 SD below the mean	>4 SD below the mean	>4 SD below the mean
Language concepts and academic skills	Difficulties in acquisition and comprehension	Limited to basic skills	Very limited, may have motor impairments	Extremely limited, may have motor and sensory impairments
Basic self-care, domestic and practical activities	Usually achieved	Achieved by some	Typically require daily support, may acquire with intensive training	Require daily support
Independent living and employment	Achieved with appropriate support	Considerable and consistent support	Supervised environment	Supervised environment

Table 1.10: Classification of intellectual disability based on severity

Abbreviations: SD, standard deviation.

The prevalence of intellectual disability is estimated at ~1% of the global population [200] and contributes a significant healthcare burden worldwide [201]. Most individuals with intellectual disability are identified with developmental delay in early childhood and severity is highly variable. Intellectual disability is clinically heterogeneous, it can occur in isolation (non-syndromic), alongside congenital malformations or multisystemic disorders (syndromic) or other neurological or neurodevelopmental features, for example epilepsy, neuromuscular features or autism. Both environmental and genetic factors contribute to intellectual disability aetiology, although the exact genetic contribution is unknown, as genetic causes may not have been identified in many (especially older) individuals with intellectual disability. The majority of severe to profound intellectual disability is thought to have a monogenic origin [202].

1.4.2 Developmental delay

Developmental delay refers to when an infant or child does not reach a developmental milestone within a defined time period. Developmental domains include: fine motor, gross motor, speech and language, cognition, personal/social and activities of daily living [203]. Developmental delay can be specific, referring to a single developmental domain, or global affecting several developmental areas. The term global developmental delay refers to children under the age of five years and is defined as

delay in two or more developmental domains. Alike intellectual disability there are standardised tests used to assess the degree of developmental delay in a child, including the Griffiths III Developmental Scale, Schedule of Growing Skills and the Bayley III Developmental Assessment, with significant delay defined as test performance over two standard deviations below the mean for age [204]. This is further classified into [203]:

- Mild: functional age <33% below chronological age

- Moderate: functional age 33-66% of chronological age

- Severe: functional age <66% of chronological age

The prevalence of global developmental delay is estimated at 1-3% of children [203]. Although many children with global developmental delay will subsequently have intellectual disability as older children and adults, this is not inevitable and a number of children will have normal cognitive functioning in adulthood.

There is huge overlap between the genetic aetiology of developmental delay, intellectual disability, and other related disorders like autism, epilepsy, schizophrenia, neuromuscular disorders and other neurological conditions, often with high co-morbidity. These conditions have been identified to share a number of pathomolecular mechanisms and signalling pathways [65]. For example, rare variants in key molecules involved in synaptic functioning, including proteins in the neurexin-neuroligin complex [205] (for example *NRXN1*, *CNTNAP2*, *NLGN1*, *NLGN3* and *LRRTM1*) and interacting proteins (for example (*SHANK1-3*, *SYNGAP1*, *DLGAP2*, *FOXP1*, *GRIN2B*, *SCN1A* and *LAMC3*) have been associated with intellectual disability, seizures, autistic spectrum disorder (ASD) and schizophrenia [206].

1.5 Neuroimaging

Individuals affected by neurodevelopmental disorders may undergo neuroimaging to identify brain abnormalities and/or malformations that may suggest a specific cause for their condition, and may assist in confirming a diagnosis [207, 208]. This can include features suggestive of environmental insults (for example hypoxia, trauma, teratogens or infection), developmental brain malformations (for example neuronal

migration disorders) or localised abnormalities (for example cerebellar and corpus callosal defects). Both computed tomography (CT) and magnetic resonance imaging (MRI) show gross morphological changes in the brain and ventricular system. MRI is usually the preferred imaging choice due to greater sensitivity with no radiation exposure, and particularly useful for detailed assessment of the developing brain in infants and children including microstructural changes, grey and white matter maturation and specific white matter pathways [209]. The posterior fossa is also more easily visualised with MRI as there are no beam-hardening artefacts.

1.6 Genetic studies in genetically isolated communities

Genetically distinctive communities may be defined by their ethnic origin, cultural or religious beliefs, and geographical location. Ten Kate *et al* (2010) define community genomics as: *"the art and science of the responsible and realistic application of health and disease-related genetics and genomics knowledge and technologies in human populations and communities to the benefit of individuals therein. Community genetics is multi-, inter- and transdisciplinary and aims to maximize benefits while minimizing the risk of harm, respecting the autonomy of individuals and ensuring equity."* [210]

Our Rare Disease Genomics research group has well established connections with healthcare professionals and researchers providing care for a number of communities internationally. The cardinal ethos of our group is to conduct translational genomic research studies that provide improvements in diagnosis and healthcare to underserved populations, which in turn provides similar benefits for other individuals affected by those genetic disorders globally. This involves the sharing of advancements in knowledge and understanding of rare (often novel) genetic conditions, the development of diagnostic strategies and clinical management guidelines, and potentially where possible the exploration and development of novel therapies. Our research findings are diagnostically validated and fed back to local healthcare providers to benefit patients, families and the wider community through accurate genetic counselling, clinical screening and management. This thesis focusses on inherited disorders of brain growth and development identified in affected individuals in families from Amish and Pakistani communities.

1.6.1 The Amish and Anabaptist communities

The Amish are a group of rural-living communities of Swiss-German ancestry that follow the Anabaptist Christian church [211]. Anabaptist means 'rebaptisers', individuals who favour adult baptism in place of infant baptism allowing individuals to confess their faith and make their own decision about membership of the church, and who advocated to separate the church from the state. Anabaptist groups were persecuted due to these beliefs throughout Europe with many fleeing to the Swiss-German mountains.

The Amish originated in Switzerland in 1693 following a schism within a group of Swiss and Alsatian Anabaptists [211]. They are known for their simple living, traditional plain dress and separation from modern technologies. One of the church group leaders, Bishop Jakob Amman whose followers became known as 'Amman-ish' (and later Amish), made a number of reforms to church practices including linking social practices to church discipline. This included enforcing a practice of shunning (or 'Meidung') whereby community members should avoid interactions with an individual if they do not follow the church order, which may ultimately result in excommunication. This is intended to encourage the individual to repent for their sinful conduct and seek forgiveness in order to be welcomed back into the church. Amman also encouraged the wearing of traditional 'plain' dress in a uniform manner and forbade beard trimming, along with other 'fashionable' changes in appearance. Other Anabaptist communities did not agree with these principles, which led to Amman's followers splitting from the wider group.

Fleeing religious persecution, a number of Amish individuals migrated to the "New-World" in two waves in the 18th and 19th centuries. Around 500 Amish families migrated to the US and settled in the county of Pennsylvania between 1736 and 1770. A second wave of ~3000 Amish immigrants travelled to Pennsylvania, Ohio, Indiana, Iowa and Illinois between 1818 and 1860. Travel to the "New-World" in the USA was attractive presenting economic opportunities with a chance to own their own land and the promise of religious freedom, in addition to avoiding military service being introduced across much of Europe. No Amish settlements remain in Europe today, the last disappeared in 1937.

Today the Amish live across 31 USA and four Canadian states, and two new settlements in Argentina and Bolivia have recently been established (Figure 1.5) [212]. The Amish are one of the fastest growing populations in the world, doubling approximately every 20 years, and the population in North America was estimated to be 361,475 in June 2021 [212]. Most Amish couples have on average 6-7 children, although in previous generations this number was greater. Amish people share general common beliefs and principles that include adult baptism, a rural life separated from modern communities, use of horse and buggy transportation, speaking German-originating dialect, non-violence, church services within homes rather than formal religious buildings, only basic education until eighth grade, traditional style of dress and very limited use of technology [213]. There are over 40 different Amish affiliations, consisting of groups of church districts each with slightly different views on religious and social practices [213].
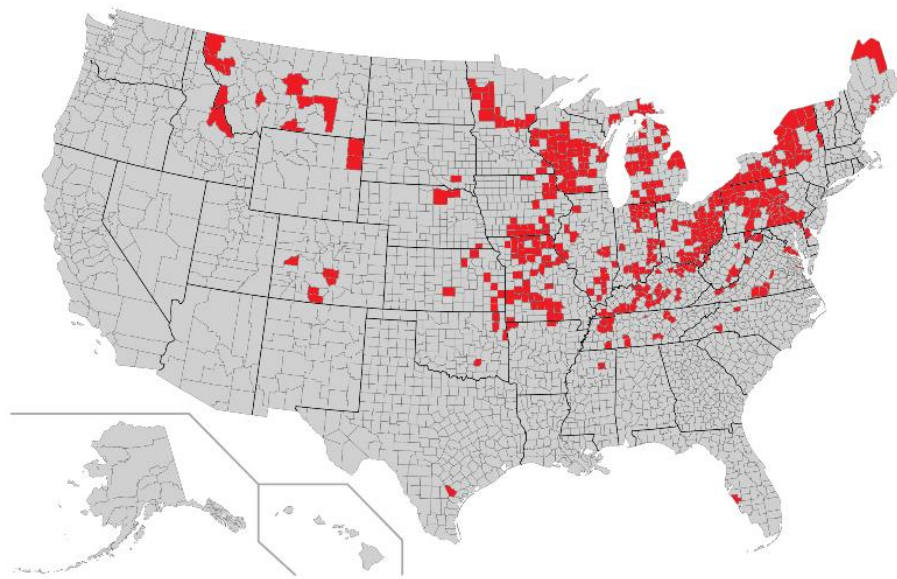
Figure 1.5: Map showing the distribution of Amish populations within the United States of America (USA) states and counties in 2020

Created using data from the Young Center for Anabaptist and Pietist Studies [212]. Reproduced from S. Braman 2021 [214] (permission granted from Steven Braman, Appendix 7).

Amish genetic distinctiveness

The relatively small number of individuals who migrated to the USA to form the current Amish comprises an excellent example of an 'ancestral bottleneck event' (Figure 1.6). The original Amish settlers were from related family groups, the gene pool from which modern day Amish communities have derived was significantly limited. Additionally very few outsiders join (or have joined) the Amish and there is an average retention rate of Amish children who decide to join the church as adults of >85% [212]. Individuals within the Amish community tend to marry within the same or similar church affiliation, leading to endogamous populations in which married couples tend to be distantly related, rather than closely consanguineous. There is also a degree of geographical isolation as the Amish are relatively immobile due to religious constraints on transportation. Thus, the resulting genetic bottleneck and subsequent population expansion has led to altered frequencies of specific allelic variants which may be rare (or absent) in other populations, which may be stably maintained and present at varied frequencies in distinct Amish communities. If altering the function of a gene or genetic element, this may result in an increased incidence of specific (typically autosomal recessive) disorders. Conversely however, other inherited disorders that are common in other populations may be seen very infrequently (or are absent) in the Amish due to this restricted gene pool [213].

Large population with genetic diversity

Bottleneck (e.g. migration, isolation) fewer individuals

Resulting population with less genetic diversity (e.g. Amish)
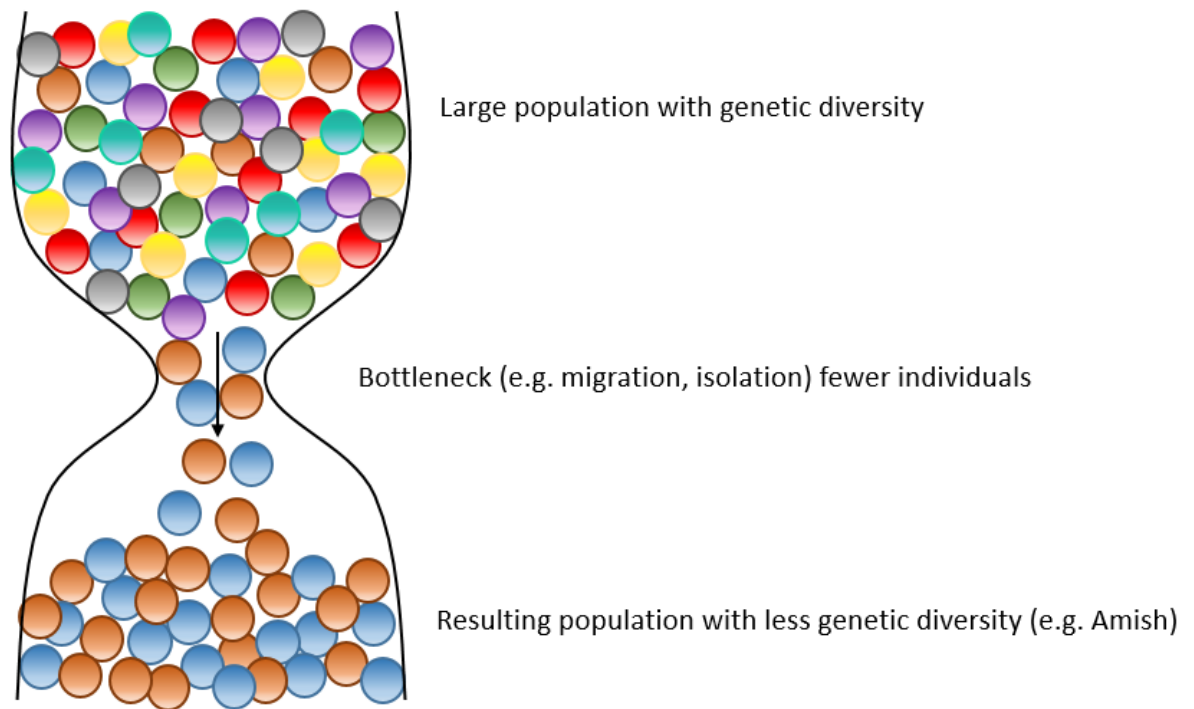
Figure 1.6: The ancestral bottleneck resulting in reduced genetic diversity of derivative populations

Many otherwise rare genetic variants present at increased allele frequency in the Amish typically arose once in the history of the community in a single founder individual (the 'founder effect'). This allelic enrichment allows the clinical relevance of a particular variant to be interpreted, whether it be benign and of no clinical relevance, or pathogenic and associated with genetic disease. Additionally, the founder gene variant nature of disease enables powerful genetic studies and approaches to discern novel genetic causes of disease. Historically this has involved 'autozygosity mapping' which may be employed to identify homozygous genomic regions common to affected individuals which are 'identical by descent', and are candidates to harbour pathogenic variants [215] (Figure 1.7). Such homozygous genomic regions were previously identified using genome-wide single nucleotide polymorphism (SNP) genotyping, although with the advent of cost-effective WES and WGS, this approach is now seldom utilised [216]. In individuals with closely related parents (for example first cousins) numerous large blocks of autozygosity are likely to be present, whereas individuals with more distantly related parents (such as in the Amish) typically have smaller and fewer blocks of autozygosity (Figure 1.7). Autozygosity mapping can be combined with WES or WGS to rapidly identify potential pathogenic variants within autozygous regions [215] including CNVs [217], and has been commonly utilised for discovery of new disease genes and proving disease causality [218, 219], including within this thesis (Chapters 3, 4 and 5).

As the Amish keep extensive birth and genealogical records, it is possible to trace ancestors back over many generations, which enables the construction of expansive family pedigrees. The Amish also tend to have large family sizes (as described above) often with multiple affected and unaffected individuals with a specific genetic condition, as well as standards of healthcare with availability of well documented medical records. Along with the genetic founder effect basis of many disorders in the community, these factors greatly empower genetic studies and facilitate the discovery of the causes and detailed clinical features of inherited conditions that would be difficult in other populations. In addition, due to the uniform standards of living and socioeconomic factors, this greatly reduces the impact of environmental variability on genetic disease expression that can be difficult to disentangle in more diverse populations.
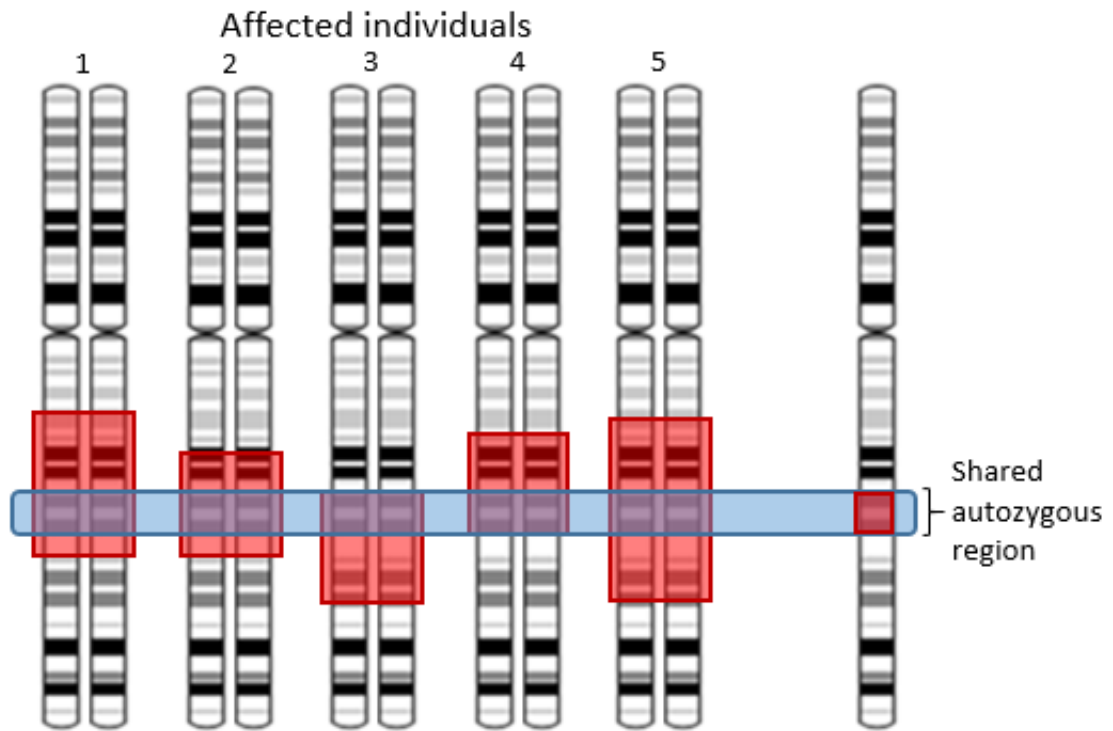
Figure 1.7: Autozygosity mapping

Autozygous regions in five affected individuals (red box) with the shared region of autozygosity indicated by the blue box and indicating the likely location of the disease-causing variant.

The Amish often have differing views about healthcare, some distrust modern medicine and will opt out of mainstream medical care. Different groups will have differing opinions on the types of medical treatment that are acceptable [213]. Few Amish people have private health insurance due to religious reasons and usually church groups will have funds to help support medical bill payments for sick individuals from the community. For these reasons, Amish patients often may not have routine investigations, for example antenatal scans or brain imaging, and post-mortem examinations are often not routinely performed, although this will depend on the beliefs or wealth of a family or their church group.

The Windows of Hope Project

Despite separation from the modern world and a lack of acceptance of many modern technologies and healthcare, the Amish are often very willing to be involved in genetic studies and research. The first geneticist to study inherited disorders in the Amish was Victor McKusick [220], who in the 1960's determined there were two distinct common types of autosomal recessive dwarfism in the community: Ellis-van Creveld syndrome and cartilage hair hypoplasia [221]. This led to further studies by McKusick and colleagues, who published over 30 papers describing inherited disorders in the Amish over the next decade [222]. McKusick's work investigating Amish genetic disease in part lead to the development and publication of Mendelian Inheritance in Man (MIM) and later the online version OMIM, which remains an invaluable core resource internationally in medical genetics today [68].

In 2000 my research supervisors Professor Andrew Crosby and Dr Emma Baple, set up a not-for-profit translational community genomics research programme called the Windows of Hope (WOH, https://wohproject.com/). The project aims to define the genetic causes of inherited conditions within the Anabaptist communities to benefit patients, families and the wider community. As well as defining many other previously described conditions which were unrecognised in the community the study has identified >25 novel genetic conditions to date. These include biallelic variants in *SNIP1* associated with a neurodevelopmental disorder with hypotonia, craniofacial abnormalities and seizures [223], *MNS1* associated with *situs inversus* and male infertility [224] and *HERC2* associated with an Angelman-like condition [225]. This research has translated into direct community benefits, through achieving a diagnosis

for many families with previously unrecognised conditions, alongside early medical intervention, education and support through understanding the clinical course of these disorders, with informed genetic counselling to support families. This has facilitated the development of accurate diagnostic testing, clinical management guidelines and targeted therapies for these conditions, which has benefitted Anabaptist patients as well as other patients worldwide who have the same conditions. Virtually all of the genetic disease genes that have been originally identified in the Amish or Anabaptist communities have subsequently been identified to cause similar diseases in other individuals worldwide, and highlights how these studies of genetically isolated populations are of great importance and relevance to patients affected by rare disease worldwide.

1.6.2 Pakistani communities

The Islamic Republic of Pakistan has a population of over 225.2 million people [226], and is the 5th most populated country in the world [227]. Population size has grown rapidly, with a six-fold increase since 1950, following improvements in the standards of living and public health [228] and is estimated to double over the next century. Pakistan is divided into 7 provinces/territories (Figure 1.8) and geographically many rural areas of Pakistani are very isolated, particularly in mountainous regions. Significant ethnic, cultural and linguistic diversity exists within the modern Pakistani population. Urdu is Pakistan's national language, although there are over 60 different spoken dialects [229]. The 2017 census reported that 96.5% of the Pakistani population is Muslim [230], with two main groups Shias and Sunnis, and several smaller groups, each further divided into several multi-layered subgroups [229]. Additionally, Pakistanis have ethnic or tribal affiliations, which form the basis for social stratification; North Pakistan comprises the major ethnic groups Punjabi, Pathan (Pashtun), Kashmiri, Kalash, Burusho and Balti, while Southern regions include Sindhi, Saraiki, Parsi, Makrani Baloch, Makrani Negroid, Muhajirs, Brahui and Blaloch. Other groups (Ismaili, Memon and Bokra) originate from Central Asia and the Middle East, with numerous smaller groups [231].

Figure 1.8: Map of the provinces and territories of Pakistan, and the location of Azad Kashmir

Pakistan is divided into four provinces (Khyber Pakhtunkhwa (KPK), Punjab, Sindh and Balochistan), one federal territory (Islamabad Capital Territory) and two autonomous territories (Gilgit-Baltistan and Azad Jammu & Kashmir). Modified from Bittles, A.H. and Small N.A. (2015) [232] (permission granted from Cambridge University Press, Appendix 7).

Consanguineous marriages are practiced by ~20% of the world population, although this is estimated to be much higher in Pakistan, with endogamous unions in over 50% of marriages [229, 233, 234], often between first and second cousins. Pakistani society values similarities in social group identity including ethnicity, religion and tribal affiliation [229]. *Biraderi*, *zat* and *quom,* meaning 'brotherhood', refer to endogamous communities in Pakistan, and imply membership of a closed patrilineal kinship group that transcends geographical boundaries and are indicators of usually immutable social status/identity acquired by birth, determined by family background and occupation, with often strict requirements for endogamy, although this depends on the specific group [232]. The elevated level of endogamy, alongside an average birth rate of 3.3 children per woman [226], has resulted in a high prevalence of autosomal recessive genetic disorders. Pakistan has one of the highest rate of genetic disorders in the world [235], including neurodevelopmental disorders, with an average of 6.2 cases of mild intellectual disability and 1.1 cases of severe intellectual disability per 100 live births [236].

In terms of a genetic perspective, different *biraderi* groups even within the same region will have distinctly different genomic profiles [237] due to the high levels of endogamy, but also the strong socio-ethnic divides between groups, and often geographical isolation in many rural areas. As with the Amish this derives from the accumulation of regional founder variants within each group (Figure 1.6), associated with specific genetic disorders [238]. Knowledge of these specific diseases and their underlying genetic basis is invaluable for local healthcare providers for these communities to aid targeted and cost-effective genetic testing, accurate diagnosis for affected individuals and their families, informed genetic counselling, and to develop clinical care pathways and policies. Development of genomic databases are ongoing in Pakistan to correlate specific genetic variants with clinical phenotype, geographical location, ethnicity or tribal affiliation (https://pakmutation.kust.edu.pk/) [239], aiming to assist clinicians, geneticists and researchers in providing personalised healthcare by providing screening for population specific variants and accurate genetic counselling.

1.7 Management of neurodevelopmental disorders

Strategies for therapeutic discovery

Identifying pathogenic variants within genes associated with neurodevelopmental disorders allows an accurate diagnosis to be achieved, with improved clinical management for the affected individual and their family, as well as identifying the underlying genetic and pathomolecular basis that may provide therapeutic targets. Recent advances in next generation sequencing (NGS) technologies of WES and WGS have led to rapid progress in the discovery of the genetic causes of many neurodevelopmental disorders. For intellectual disability, 1139 causative genes have been confirmed (PanelApp green genes, Genomics England (https://panelapp.genomicsengland.co.uk/) accessed March 2022), with many more proposed intellectual disability genes (SysNDD database lists 2841 genes (accessed June 2022 [240]). These increases in knowledge and understanding are resulting in increased diagnostic yield of genetic tests. WES or WGS can identify causal genetic variants in around 50% of cases [241-243]. This is translating into direct benefits for patients and their families, with improved likelihood of a genetic diagnosis, availability of prenatal and carrier testing, screening for possible complications and potential therapeutic advances with available treatments in some cases. Overall NGS has the potential to transform the diagnosis and molecular understanding of inherited disorders to facilitate the provision of personalised medicine and in the UK is currently being integrated into NHS services following the success of the 100,000 genomes project.

Within this thesis a combined approach of variant detection with WES and autozygosity mapping to identify regions of shared homozygosity between affected individuals was used as a powerful tool to identify novel causes of neurodevelopmental disorders in endogamous communities.

1.8 Project aims

The overarching objective of this PhD involves the clinical and genetic delineation, and functional characterisation, of inherited disorders of brain growth and development identified within Amish and Pakistani communities. To this end, I pursued the following aims:

1. to perform detailed clinical phenotyping, and comprehensive genetic and functional studies to confirm that biallelic pathogenic variants in a novel candidate disease gene, *TRAPPC10,* cause a new microcephalic neurodevelopmental "TRAPPopathy" disorder,

2. to undertake in depth clinical and genomic studies to expand knowledge on the megalencephalic neurodevelopmental disorder *KPTN*-related syndrome. This work will define the phenotypic and genetic spectrum of the disorder for comparison with a *kptn* knockout mouse model. Together, this will aid the development of clinical management guidelines and inform future studies on potential therapeutic interventions,

3. to perform clinical phenotyping in combination with molecular studies in Amish and Pakistani families with individuals affected by undiagnosed forms of inherited disorders of brain development, and where appropriate explore GeneMatcher and our international collaborative network to identify additional cases and further characterise and consolidate the genetic and clinical spectrum of each condition identified.

4. to translate research findings from my study into patient, family and community healthcare benefits through improved diagnostic and prognostic information, genetic counselling, clinical management and ultimately targeted therapies.

2 MATERIALS AND METHODS

2.1 Reagents, buffers and stock solutions

Reagents, buffers and stock solutions used in this study are listed in Tables 2.1 and 2.2. Laboratory consumables were purchased from Rainin, Starlabs and Alpha Laboratories.

Reagents	Manufacturer
Agarose (molecular grade)	Fisher Scientific
Ethidium bromide	Fisher Scientific
Redsafe™ Nucleic Acid Staining Solution (20,000x)	Intron Biotechnology
GeneRuler™ 1kb Plus DNA ladder	Thermo Fisher Scientific
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
Deoxyribonucleoside triphosphate (dNTP) solution mix	Solis BioDyne
DreamTaq™ DNA polymerase	Thermo Fisher Scientific
DreamTaq™ green buffer (10x)	Thermo Fisher Scientific
DreamTaq™ Green PCR Master Mix (2x)	Thermo Fisher Scientific
Exonuclease I	New England Biolabs
Shrimp alkaline phosphatase (rSAP)	New England Biolabs
Ethanol	Fisher Scientific
Protease inhibitor cocktail tablets (EDTA-free)	
Ampicillin	
Kanamycin	
Penicillin 100U/ml)	
Streptomycin 100U/ml	
Trypsin	
Beta-mercaptoethanol	
Milk powder	
Tween20 (detergent)	
Neomycin (G418) 735µg/mg	Multicell
JetPrime DNA/siRNA transfection reagent	Polyplus
Glycerol 20%	
Methanol	
Amersham ECL™ Start Western Blotting Detection Reagent	Cytiva
Mini-Protean TGX™ precast gels	Bio-Rad
Bovine serum albumin (BSA)	Bio-Rad
Colour protein standard	New England BioLabs (NEB)
Loading dye (6x)	
2.1 restriction enzyme buffer	
3.1 restriction enzyme buffer	
Cutsmart buffer	
Bsrgl restriction enzyme	
EcorV restriction enzyme	
Sac1-HF restriction enzyme	
BP Clonase™ II enzyme mix	Invitrogen
LR Clonase™ II enzyme mix	Invitrogen
Proteinase K	
C-terminus red fluorescent protein (mRFP-C1)	Addgene #54764

Table 2.1: Reagents used in this study and manufacturer

Buffer/Stock solution	Composition
Lithium acetate borate (LAB) buffer	For 1 litre 50x stock solution: 51g lithium acetate dehydrate, 31g boric acid in ddH$_2$O to final volume
Tris borate EDTA (TBE) buffer	For 1 litre 5x stock solution: 54g Tris base, 27.5g boric acid, 20ml 0.5M EDTA in ddH$_2$O to final volume
ExoSAP enzyme mix	For 1ml: 2.5µl Exo I (20,000U/ml), 25µl rSAP (1000U/ml) in ddH$_2$O to final volume
Luria Bertani (LB) medium	For 1 litre: 5g yeast extract, 10g tryptone, 10g NaCl in ddH$_2$O to final volume, then autoclaved
LB agar	As above but add 15g agar, then autoclave
Dulbecco's modified eagle medium (DMEM)	Cell culture medium containing glucose, amino acids, vitamins and minerals.
Iscove's modification of Dulbecco's modified eagle medium (IMDM)	Contains additional amino acids, vitamins, selenium and HEPES buffer, and is formulated with potassium nitrate in place of ferric nitrate.
Fetal bovine serum (FBS)	The liquid fraction (serum) remaining after blood drawn from bovine fetus coagulates and is centrifuged
Dulbecco's phosphate buffer solution (PBS)	Dulbecco's PBS without calcium and magnesium.
Lysis buffer	For 20ml stock: 1ml 1M Tris (pH 7.2), 1.5ml 2M NaCl, 20µl 0.5M EDTA, 20µl 1M DTT, 2µl triton 10%, 2 X protease inhibitor tablets in ddH$_2$O to final volume
SDS Sample buffer (4x)	For 1 litre: 200ml 1M Tris-HCl, 400ml 1M DTT, 80g SDS, 4g bromophenol blue, 320ml glycerol in ddH$_2$O to final volume
Resolving buffer (4x)	For 1 litre: 181.7g Tris base, 4g SDS, adjusted to pH8.8 using HCl, in ddH$_2$O to final volume
Stacking buffer (4x)	For 1 litre: 60.6g Tris base, 4g SDS, adjusted to pH6.8 using HCl, in ddH$_2$O to final volume
Running buffer (10x with 1% SDS)	For 2 litres: 60.4g 250mM Tris, 376g 2.5M glycine in ddH$_2$O to final volume
Transfer buffer	For 1 litre: 100ml 10x transfer buffer, 200ml methanol, 700ml ddH$_2$O
Phosphate buffer solution (PBS)	For 1 litre: 80g NaCl, 2g KCl, 14.4g Na$_2$HPO$_4$, 2.4g KH$_2$PO$_4$ in ddH$_2$O to final volume
PBST	For 500ml: 50ml 10X PBS, 450ml ddH$_2$O, 250µl tween 20
BioRad protein assay dye reagent solution	Mix 1 part dye reagent with 4 parts ddH$_2$O. BioRad

Table 2.2: Buffers and stock solutions used in this study, including their composition

2.2 Subjects

2.2.1 Ethical approval

Ethical approval for this study was granted by the following Ethical approval committees:

- Institutional Review Board, International Islamic University, Islamabad, Pakistan (IRB number IIU/D(BS)/FBAS/-2020-1309)

- Institutional Review Board of the Office for Responsible Conduct of Research, University of Arizona, Arizona, USA (IRB number 10-0050-01)

- Akron Children's Hospital, Ohio, USA (IRB number 986876-3)

- University of Exeter Research Ethics Committee, Exeter, UK (approved IRBs above)

All studies were carried out in accordance with the principles of the Declaration of Helsinki. All participants of the study, or their legal guardian(s), agreed to inclusion within the study and written informed consent was obtained prior to collecting any clinical details or information from medical records, photographic images or blood/buccal samples for genetic testing.

2.2.2 Recruitment and phenotyping of affected individuals

Recruitment of families from Amish and Pakistani communities was conducted by researchers and clinicians within well-established collaborative clinical and academic institutes. Other families were identified internationally through their responsible treating clinician via submission to GeneMatcher [244], pre-established collaboration with research groups, the Deciphering Developmental Disorders (DDD) project, the 100,000 genomes project or direct contact from local clinicians or families.

Phenotypic information was obtained through the clinical care provider. A full medical and developmental history and systemic examination for the purposes of the research study was obtained for all individuals recruited and photographic records and/or videos of pertinent examination findings were taken with informed consent. Where possible detailed phenotype information was recorded using a targeted questionnaire to identify specific clinical features. Height, weight and OFC were measured in cm using standard methods. Standard deviations and Z-scores were calculated using the LMS method to access growth references (UK-WHO, British 1990 and WHO 2006) using a Microsoft Excel add-in [245]. Copies of relevant medical records, investigations (for example neuroimaging, EEG studies, audiograms) and educational reports were requested using the appropriate information release form. Whenever possible original neuroimaging was requested for review locally by Dr Lucy McGavin, Consultant

Neuroradiologist, University Hospitals Plymouth NHS Trust, Plymouth. Detailed information about the wider family and other possibly affected individuals was recorded in the form of a family pedigree. For Amish families, pedigrees were constructed using detailed birth and family records from the Swiss Anabaptist Geneological Association (SAGA) database (www.saga-omii.org).

Some of the families in this study live in remote rural locations with limited medical infrastructure. This proved challenging for local clinicians/researchers assessing these families, as often there may be no or very limited medical records or investigations due to access or associated costs. Local clinicians and scientists were supported by myself (clinical genetics trainee) in the collection and review of the highest quality clinical phenotypic data possible.

2.2.2.1 Psychometric assessment

Detailed psychometric assessments were performed using the Wechsler Intelligence Scale for Children 4th Edition (WISC-IV) [198] (6-16 years) or the Wechsler Adult Intelligence Scale 4th Edition (WAIS-IV) [199] (over 16 years), which assesses cognitive performance in four domains including verbal comprehension (VCI), perceptual reasoning (PRI), processing speed (PSI) and working memory (WMI) indices, and can be combined to generate a full-scale intelligence quotient (FSIQ) score. Immediate memory recall was assessed using the story memory test and list learning test from the developmental NEuroPSYchological Assessment- 2nd Edition (NEPSY-II) [246] (5-16 years) or the Repeatable Battery for Assessment of Neurological Status- 2nd Edition (RBANS-Update) [247] (over 12 years). The WISC-IV and WAIS-IV tests have normalised reference scores, with mean (SD) of 100 (15). The NEPSY-II and RBANS-Update tests generate scaled scores, with mean (SD) of 10 (3). The assessments were carried out by my supervisor Dr Emma Baple who was trained and supervised by Dr James Tonks, Consultant Paediatric Neuropsychologist and Consultant Clinical Psychologist, University of Exeter Medical School. Analysis of the data was undertaken by myself with support from Dr Tonks to calculate standardised test scores that were converted to Z-scores for direct comparison between the tests.

2.3 Sample acquisition and storage

Study participants agreed to sample collection following written informed consent. Peripheral venous blood samples were collected in EDTA Vacutainer™ tubes, or buccal cells were collected using the ORAcollect® kit (DNA Genotek) for DNA extraction. For RNA studies and creation of lymphoblastoid cell lines in some participants peripheral venous blood was collected in PAXgene™ blood RNA and ACD Vacutainer™ tubes respectively.

On arrival blood samples were stored at -20°C and buccal samples were stored at room temperature before DNA extraction was performed.

This study was conducted in accordance with the Human Tissue Authority (HTA) Codes of Practice and Standards for Human Tissue and Research (code E). 'Relevant material' included under the Human Tissue Act 2004 is defined as any material *'which consists of or includes human cells'* and includes blood samples. All blood, buccal and DNA extractions used in this study were stored in HTA-licenced premises and research conducted in accordance to the Human Tissue Act 2004.

Data management

On collection of a sample, each individual is given a participant ID, and each sample a sample ID. Clinical and molecular information is recorded in a password-protected database stored on University of Exeter servers. Management of all participant data is compliant with ethical approvals and General Data Protection Regulation (GDPR) legislation.

2.4 Molecular genetic methods

2.4.1 DNA extraction and quantification

DNA extraction from whole blood

Genomic lymphocyte DNA extraction was performed using the ReliaPrep™ Blood gDNA Miniprep System (Promega) on peripheral venous whole blood samples, as summarised below. For all stages filter pipette tips were used to prevent contamination of samples. Blood samples were thoroughly mixed by hand for at least 2 minutes. When frozen, blood samples were completely thawed and mixed for 10 minutes prior to DNA extraction. 200μl whole blood was added to a 1.5ml microcentrifuge tube

containing 20μl of Proteinase K and briefly mixed. Then 200μl of cell lysis buffer was added to the microcentrifuge tube, the contents vortexed for 10-20 seconds and then incubated on a heating block at 56°C for 10 mins. The tube was then removed from the heating block and 250μl of binding buffer was added, followed by vortexing for 10-20 seconds to mix the contents. The lysate colour was confirmed to be dark green according to the protocol before transferring the contents to a ReliaPrep™ binding column placed within an empty collection tube. This tube was then microcentrifuged at maximum speed of 13,000rpm for 1 min and checked to ensure all the lysate had passed through the column membrane, if any lysate remained visible on top of the membrane it was centrifuged for a further minute. The collection tube was then removed and the liquid discarded as hazardous waste and the binding column placed within a fresh collection tube. 500μl of column wash solution was added to the column and centrifuged at maximum speed for 3 minutes before discarding the flowthrough. This wash step was repeated to complete a total of three washes. The binding column was then placed in a clean, autoclaved 1.5ml microcentrifuge tube and 200μl of nuclease free water added. The column was then centrifuged at maximum speed for 1 min to elute the DNA.

DNA extraction from buccal samples

Genomic DNA extraction from buccal swabs was performed using the *Xtreme* DNA kit (Isohelix). The buccal swab heads contained within the sample tubes were incubated in a water bath at 50°C for 1 hour. 500μl of lysis buffer was added to the sample tube to cover the swab head and vortexed, followed by addition of 20μl Proteinase K and vortexed to mix. The swab sample was further incubated at 60°C for 1 hour to lyse the sample. 750μl of column binding buffer was added and thoroughly mixed by vortexing for 30 seconds, followed by 1.25ml ethanol and vortexed to mix. An Xtreme DNA column was placed onto a collection tube and 700μl of sample added to the column without touching the rim. This was microcentrifuged at maximum speed of 13,000rpm for 1 minute and flow through discarded, then this step was repeated until all the sample had been passed through the column membrane. 750μl of wash buffer solution was added to the column and centrifuged at maximum speed for 1 minute, discarding the flow through. This step was repeated by adding another 750μl of wash buffer solution for a total of two washes. The column was placed in a clean collection tube and centrifuged at maximum speed for 3 minutes to remove any excess ethanol. The

column was then placed in a clean 1.5ml microcentrifuge tube and 100µl of elution buffer preheated to 70°C was added to the centre of the membrane and left to stand for 3 minutes. The DNA was then eluted by centrifuging the tube at maximum speed for 1 minute.

DNA quantification and storage

Concentration and purity of the extracted DNA within the eluate was measured using a spectrophotometer (NanoDrop™ 2000, Thermo Fisher Scientific), by measuring the absorption of UV light at 260nm (A260) of a 1µl undiluted sample, this is the peak absorbance wavelength of nucleic acids. The NanoDrop™ software uses a modified Beer-Lambert equation to automatically calculate the nucleic acid concentration in ng/µl. DNA purity was simultaneously assessed by measuring absorption at 280nm (A280) and 230nm (A230), the peak absorbance wavelengths for proteins/phenolic compounds and organic compounds respectively. The A260/A280 and A260/A230 ratios were then calculated. An A260/A280 ratio of ~1.8 indicates a 'pure' nucleic acid sample, with lower values often indicating the presence of contamination with proteins or other molecules. The A260/A230 ratio is often higher with 'pure' nucleic acid samples between 1.8-2.2, with lower values often indicating the presence of contaminants co-purified with the DNA. Once quantified, extracted DNA samples were then labelled and stored at 4°C.

2.4.2 RNA extraction

RNA was extracted from lymphocytes in peripheral venous whole blood collected in PAXgene™ collection tubes. RNA extractions for this study were performed with assistance from Joseph Leslie (Research Assistant, University of Exeter). To prevent contamination of samples and equipment, prior to commencing RNA extraction the lab bench, rack and pipettes were cleaned with RNase decontamination solution (RNaseZap™) and filter pipette tips were used at all stages. RNA samples were kept on ice to reduce degradation.

RNA extraction was performed using the PAXgene® Blood mRNA Kit from Qiagen following the automated protocol for use with a QIAcube robot (Qiagen). PAXgene™ sample blood tubes were centrifuged using a swing out centrifuge at 3000rpm for 10 minutes. The resulting supernatant was removed by pipetting without disturbing the pellet and 4ml of RNase-Free water added, followed by tube closure using a fresh BD

Hemogard closure. The tube was then vortexed until the pellet was resuspended and centrifuged at 300rpm for 10 minutes, then the supernatant removed by pipetting and discarded. 350µl of BR1 buffer was added and vortexed until the pellet was resuspended. The sample was then pipetted into a 2ml processing tube, with the open processing tubes then loaded into the QIAcube instrument shaker and the 'PAXgene Blood RNA Part A' protocol commenced with automated steps completed by the QIAcube robot. Once completed, the 1.5ml microcentrifuge tubes containing the purified RNA are removed and transferred onto the QIAcube instrument shaker adapter before starting the 'PAXgene Blood RNA Part B' protocol. After completion of this program the eluted RNA samples were immediately removed and placed on ice before quantification (or stored at -70°C until required).

RNA sample concentration, integrity and quality were measured with the RNAScreenTape Assay (Agilent) which uses microcapillary electrophoresis to analyse nucleic acids. In a 0.2ml reaction tube strip 5µl RNA sample buffer and 1µl RNA ladder were added to the A1 position. Then for each RNA sample 1µl was added to 5µl RNA sample buffer and the tubes sealed. Samples were then vortexed at 200rpm for 1 minute followed by microcentrifugation at maximum speed for 1 minute. Samples were heated to 72°C for 3 minutes and then placed on ice for 2 minutes. The tubes were then centrifuged at maximum speed for a further 1 minute before being loaded into the Agilent 4200 TapeStation instrument with the ladder at position A1 in the tube strip holder. The reaction tube strip cap was removed and the TapeStation analysis commenced. RNA sample concentration and RNA integrity number (RINe) values were calculated [248]; the lowest RIN score of 1 indicates the most degraded RNA profile and the highest score of 10 indicates highly intact RNA. A RIN score over 8 was preferable, although samples with lower scores were used for some subsequent methods. Once quantified, extracted RNA samples were then labelled and stored at -70°C.

2.4.3 Gel DNA extraction

Sequencing of DNA from a PCR product run on an agarose gel was carried out when there was more than one clear PCR product band and optimisation of the PCR was difficult. The band corresponding to the expected PCR product size was cut out of the gel using a direct UV imager and placed into a 1.5ml microcentrifuge tube. The sample

was then weighed to ensure it was less than 300mg. The Gel/PCR DNA Fragments Kit (Geneaid) was then used to extract DNA from the gel. 500µl of DF buffer was added to the gel, vortexed and then incubated at 60°C for 10-15 minutes, inverting the tube every 2-3 minutes to completely dissolve the gel before allowing the mixture to cool to room temperature. A DF column was placed in a 2ml collection tube and 800µl of sample mixture was transferred to the DF column before centrifuging at 14,000rpm for 30 seconds. Flow through was discarded and the DF column placed back in the collection tube prior to adding 400µl of W1 buffer and centrifuging at 14,000rpm for 30 seconds. Flow through was discarded and 600µl of wash buffer was added in the column and allowed to stand for 1 minute before centrifuging at 14,000rpm for 30 seconds. Flow through was discarded and the column and tube centrifuged at 14,000rpm for 3 minutes to dry the column matrix. The dried DF column was then transferred to a new 1.5ml microcentrifuge tube and 20µl of elution buffer was added into the centre of the column matrix and left to stand for at least 2 minutes to ensure the elution buffer was completely absorbed. The purified DNA was eluted by centrifuging at 14,000rpm for 2 minutes.

2.4.4 Polymerase chain reaction (PCR)

Oligonucleotide primers for dideoxy sequencing were designed using Primer 3 software (https://primer3.ut.ee/). Reference genomic sequences were accessed from the UCSC Genome Browser (https://genome.ucsc.edu/). Primers were designed using the following criteria:

- Primer length between 20-26 base pairs

- Melting temperatures similar for each of the primer pairs, ideally within 5°C of each other

- Guanine-cytosine (GC) content between 40-60% if possible

- A resulting PCR product size of 200-1000bp if possible

- Primer sequences are uniquely specific and a 100% match to the targeted region to be amplified (checked with *in silico* PCR and BLAT searches on UCSC)

- Primers sequences should not contain any common SNP's (minor allele frequency >1% in dbSNP 153), repetitive elements (interspersed repeats e.g. SINE, LINE and low complexity sequences), microsatellites (e.g. di/trinucleotide repeats) and simple repeats identified with RepeatMasker, Interrupted Rpts, Microsatellite and Simple Repeats tracks within the UCSC Genome Brower.

If it was not possible to design primers according to these criteria using Primer3, for example when regions flanking the target sequence were GC rich or highly repetitive and not suitable for automated primer design, manual primer design was sometimes required. This involved visualisation of the DNA sequence surrounding the target and selection of optimal primer locations based on the above principles.

DNA Oligonucleotide primers were manufactured by Integrated DNA Technologies (IDT) and received as a freeze-dried powder form requiring resuspension in double distilled water (ddH$_2$O) to result in a 100µM stock solution. Stock primers were stored at -20°C and a working dilution of 200µl (5µM) was made for use in polymerase chain reaction (PCR) and stored at 4°C. Details of DNA primers used in this study and PCR conditions are listed in Appendix 1 Table A1.1.

PCR was performed using a standard 10µl reaction (Table 2.3).

Reagent (concentration)	Volume
DreamTaq™ Green PCR Master Mix (2x)	5µl
DNA-free H$_2$O	3.4 µl
Forward primer (5µM)	0.4µl
Reverse primer (5µM)	0.4µl
DNA (10-30ng/µl)	0.8µl
Total	10µl

Table 2.3: Standard PCR Reaction Mixture

The DreamTaq™ Green PCR Master Mix contains Taq polymerase, a mixture of dNTPs (dATP, dCTP, dGTP and dTTP), as well as the constituents of DreamTaq™ optimised green buffer (MgCl$_2$ (20mM), a density reagent and dyes for direct loading

and visualisation of products on the agarose gel). Reagents containing DNA taq polymerase were kept on ice to optimise enzymatic function.

If the PCR reaction produced weak or no product, then the contents of the mix and amplification conditions were altered (see optimisation of amplification conditions below). In order to ensure that the desired DNA template was being amplified and not a contaminant, a DNA-free H_2O control was included in each reaction.

The PCR mixture was placed in a themocycler (Eppendorf Mastercycler® ep Gradient S) and PCR performed using a 'touchdown' PCR protocol, which improves the binding specificity of primers and helps to reduce amplification of other genomic regions [249]. This method uses an initial annealing temperature (Ta) that is greater than the predicted melting temperature (Tm) of the primers, then incrementally lowers the annealing temperature by 2°C every two cycles until the target annealing temperature is achieved (Table 2.4).

Step	Temperature	Duration	Cycles
Initial denaturation	95°C	5 mins	1
Denature	95°C	30 sec	
Anneal	Ta + 4°C	30 sec	2
Extension	72°C	30-60 sec	
Denature	95°C	30 sec	
Anneal	Ta + 2°C	30 sec	2
Extension	72°C	30-60 sec	
Denature	95°C	30 sec	
Anneal	Ta	30 sec	35
Extension	72°C	30-60 sec	
Final extension	72°C	5 mins	1

Table 2.4: Touchdown PCR protocol

The size of the PCR product determined the duration of the extension step:

- <500bp – 30 second extension

- 500 to 750bp – 45 second extension

- >750bp – 60 second extension

The melting temperature of an oligonucleotide primer is the temperature at which 50% of double-stranded DNA is changed to single stranded DNA. The annealing temperature was set a few degrees (2-4°C) lower than the lowest Tm of the primer pair, to allow both forward and reverse primers to bind to the single stranded DNA but to avoid the formation of complexes through non-specific binding. If these conditions did not result in a clean, specific PCR product with a good yield then the PCR reaction for the primer pair was optimised using a temperature gradient to determine the optimal primer annealing temperature to achieve a high yield with minimal non-specific amplification. A temperature gradient of 54-66°C was used for 12 PCR reactions for the primer pair using good quality control DNA, with temperature increasing incrementally by ~1°C across each of the 12 lanes of the thermocycler PCR block (Figure 2.1).

For some oligonucleotides it was not possible to achieve a primer product with less than 60% GC content. If the resultant PCR reaction showed non-specific binding or poor PCR product yield this may be caused by the increased hydrogen bond strength from the high proportion of GC bonds increased the difficulty in denaturing template DNA, resulting in intermolecular structures that can compete with primer annealing [250]. Dimethyl sulfoxide (DMSO) can bind to and destabilise the template DNA to promote denaturation and reduce secondary structure formation. PCR reactions with products of high GC content were supplemented with 10% DMSO.
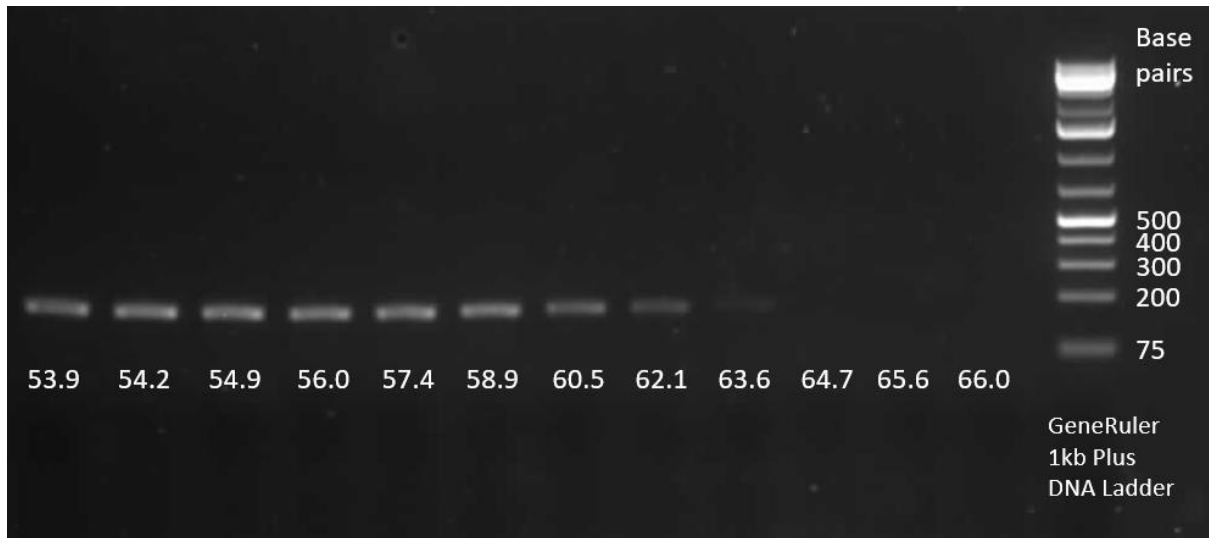
Figure 2.1: Temperature gradient PCR optimisation example

Agarose gel electrophoresis results for PCR reaction temperature gradient. The optimal annealing temperature used was 55°C as this is the lowest temperature producing a bright clear product band.

Agarose gel electrophoresis

Agarose gel electrophoresis was used to confirm the presence of a single PCR amplification product and determine the adequacy of DNA amplification. A 1% or 2% agarose gel was used for resolving smaller DNA fragments (usually <500bp) and larger fragments were resolved with ≤1% agarose gels prepared using similar methods. A 1% agarose gel was made by adding 1.0g of agarose to 100ml of 1X lithium acetate borate (LAB) buffer and heating in a microwave (700W) until the agarose completely dissolved. The solution was left to cool to 50-60°C before 5µl of 1% ethidium bromide was added and swirled to mix. The agarose solution was then poured into a casting tray with wells created by the insertion of combs, any bubbles were removed and this was left to solidify at room temperature for 30-60 minutes. Once set the gel was placed into an electrophoresis tank filled with 1x LAB buffer to sufficiently cover the surface of the gel. 5µl of a molecular weight marker (GeneRuler™ 1kb Plus DNA Ladder) was directly loaded into one well to assess the size of the PCR product and 3µl of each PCR product was directly loaded into separate wells. The samples were electrophoresed through the gel at a constant voltage of 150V for 25-30 minutes before being visualised on a UV transilluminator resulting in fluorescence of the ethidium bromide bound to the DNA bands. Images were recorded using a gel imaging system (InGenius gel system and GeneSnap image software (Syngene)).

Purification of PCR products

Prior to dideoxy sequencing of PCR products it is necessary to remove any unincorporated primers and dNTP's. This was performed using ExoSAP enzyme mix, which is a combination of exonuxlease I (Exo I) which degrades single stranded DNA by releasing deoxyribonucleoside 5 prime monophosphates from excess PCR products and residual primers, and shrimp alkaline phosphatase (SAP) which catalyses the release of 5 and 3 prime phosphate groups from extra nucleotides to result in their deactivation. 2µl of ExoSAP was added to 5µl of each PCR product and incubated at 37°C for 30 minutes for optimum enzyme activity followed by incubation at 85°C for 5 minutes to inactivate the enzymes.

Sequencing reaction

Cleaned up PCR products were then sent for dideoxy sequencing at Source BioScience (https://www.sourcebioscience.com/) using capillary electrophoresis. This

involves chain termination PCR using fluorescently labelled dideoxyribonucleotides (ddNTP's), followed by size separation by capillary gel electrophoresis and laser detection of the incorporated fluorescent tags to call the identity of each terminal ddNTP and determine the DNA sequence. Chromotogram data files were then viewed with ChromasLite (Technelysium Pty Ltd), or FinchTV (Geospiza, Inc) software.

2.4.5 Reverse transcription PCR (RT PCR)

Reverse transcription PCR (RT PCR) is a method used to indirectly amplify and determine the sequence of RNA. A complementary DNA (cDNA) copy is first created from the RNA using reverse transcriptase enzyme, then amplified and sequenced from which the RNA sequence can be inferred.

To prevent contamination of samples and equipment, prior to commencing RT PCR the lab bench, rack and pipettes were cleaned with RNase decontamination solution (RNaseZap™) and filter pipette tips were used at all stages. Extracted RNA was stored on ice until required. The EvoScript Universal cDNA Master kit (Roche) and protocol were used and two-step RT PCR was performed with the assistance of Joseph Leslie, University of Exeter. The first step of the RT PCR was cDNA synthesis from the RNA samples. Table 2.5 lists the reaction components.

Reagent (concentration)	Volume	Final concentration
PCR grade H_2O	x	-
Reaction buffer (5x)	4µl	1x
Enzyme mix (10x)	2µl	1x
Template RNA	x	2.5µg (down to 1pg)
Total volume	20µl	

Table 2.5: Reaction components for cDNA synthesis

The amount of PCR grade H_2O to be added was determined by the volume of template RNA added for a final concentration of 1pg - 2.5µg and a total volume of 20µl. All the reagents except the enzyme mix were added to a nuclease free reaction tube placed on ice, mixed and briefly centrifuged, then incubated on ice for 5 minutes to allow the

primers to anneal to the RNA. Finally, 2µl of enzyme mix was added before placing the tubes in a thermocycler for incubation following the reverse transcription protocol:

- Heat to 42°C for 15 minutes

- Heat to 85°C for 5 minutes

- Heat to 65°C for 15 minutes

- Cool to 4°C with an unlimited hold time

- Stop the reaction by placing the tube on ice

The enzyme mix has RNase H activity which removes the RNA template after cDNA synthesis, allowing PCR primers to easily bind to the cDNA. The cDNA samples were then stored at -15 to -25°C until required for PCR. The second step of the RT PCR was then to perform a PCR using the same methods as described previously to amplify the region of interest within the cDNA. This was then sent for dideoxy sequencing as described previously and the RNA sequence inferred from the results.

RT PCR primers were designed to target exonic sequences that are separated by at least one intron. Primers were designed using Primer3 following the methods as described earlier, then resuspended and diluted as previously described. Details of RT PCR primers used in this study and PCR conditions are listed in Appendix 1 Table A1.2.

2.4.6 Single nucleotide polymorphism (SNP) genotyping

Genome-wide SNP genotyping was carried out using Illumina HumanCytoSNP-12v2.1 arrays following the Infinium® HD Assay Ultra manual protocol. This assay requires 200ng of DNA (50ng/µl) per sample, and is performed over three consecutive days:

Day 1: Denature DNA samples using a 0.1N sodium hydroxide buffer, then neutralised and incubated at 37°C overnight for isothermal whole genome amplification.

Day 2: Enzymatic fragmentation of amplified DNA samples using Illumina Fragmentation solution (FMS), followed by precipitation of DNA with 100% 2-propanol and Illumina precipitation solution (PM1). The DNA was collected by centrifugation at 4°C followed by resuspension using Illumina resuspension, hybridisation and wash solution (RA1), and denaturation at 95°C for 20 minutes. Denatured DNA samples

were cooled at room temperature for 30 minutes and then 12µl sample loaded onto the BeadChip, which was then incubated at 48°C for 16-24 hours in the Illumina Hybridisation Oven.

Day 3: BeadChips were washed with Illumina Prepare BeadChip Buffer 1 (PB1), a solution containing an aliphatic amide, to remove any un-hybridised and non-specifically hybridised DNA. Labelled nucleotides were dispensed onto the BeadChip through the flow-through chambers to allow single-base extension of primers hybridised to captured DNA on the BeadChip. BeadChips were then stained using the Illumina XStain HD BeadChip process before imaging on an Illumina iScan System, which uses a laser to excite the fluorophores of the single-base extension products on the beads. A reader takes high resolution images of the BeadChip to record fluorophore light emissions from the BeadChip.

The Illumina GenomeStudio Integrated Informatics Platform analyses data from these images to determine the genotype. This data was then exported into Microsoft Excel 2016 and analysed using a macro (written by Dr Barry Chioza, University of Exeter) and visually to identify notable regions (>1Mb) with a shared haplotype. These regions were also visualised using Illumina KaryoStudio v1.4 to identify CNV's or other chromosomal structural aberrations.

Assistance and support for SNP chip genotyping and analysis was provided by Joseph Leslie and Dr Barry Chioza.

2.4.7 Next Generation sequencing (NGS)

Whole exome sequencing (WES)

Whole exome sequencing (WES) was performed by the Exeter Sequencing Service using either (i) an Illumina NextSeq 500 sequencer (University of Exeter, UK), (ii) Illumina HiSeq (Royal Devon and Exeter NHS Foundation Trust, UK) or (iii) by BGI Tech Solutions (BGI Genomics, Hong Kong) using a BGISEQ-500 platform. Targeting involved Agilent Sureselect Whole Exome v6 or Twist Human Core Exome v2. For all platforms a minimum read depth and coverage of 20x was achieved for 90-95% of the exome.

The bioinformatics pipeline consisted of the following:

- **Read alignment/mapping:** Reads from raw sequence data in FASTQ format were aligned to the reference human genome (Genome Reference Consortium human genome build 37 (GRCh37)) using Burrow-Wheeler Alignment minimal exact matches (BWA-MEM) (v0.7.17) [251].

- **Conversion to BAM file:** Output files in Sequence Alignment Map (SAM) format were then converted to (Binary Alignment Map) BAM format files for further analysis using SamFormatConverter, Picard Tools v2.15.0 [252].

- **Duplicates removed:** Read pair information was verified and duplicate reads were removed using FixMateInformation and MarkDuplicates, Picard Tools v2.15.0 [252].

- **Indel realignment:** The BAM file was then realigned for indel variants using the Genome Analysis Toolkit (GATK) v3.7.0 [253].

- **CNV calling:** CNV's were detected using either ExomeDepth [254] (uses read depth data) or SavvyCNV [255] (uses data from off-target reads).

- **Base quality score recalibration:** performed using BaseRecalibrator GATK v3.7.0 [253].

- **Variant calling:** For SNV's and small indels variants were called using HaplotypeCaller GATK v3.7.0 [253].

- **Filtered variant call format (VCF) file generation:** A VCF file was created with quality control filtering based on read depth (DP), mapping quality (MQ), relative positioning of the variant within the read and strand bias. Further filtering with a capture-kit specific list of common variants, including gnomAD frequency >2%, dbSNP 'common no known medical impact' file and a list of internal artefacts.

- **VCF annotation:** Functional annotation of variants within the filtered VCF file was undertaken using Alamut Batch (v1.10). Each variant was annotated on all available transcripts.

- **Filtered Excel file generation:** The VCF file was further filtered based on positive and negative filters or inheritance in the case of trio or parental exome (Table 2.6A,B) to generate a filtered Microsoft Excel file for analysis.

A Global filters

Positive filters (keep all):

	HGMDPro	Present (AND not retired)
OR	ClinVar	Contains "pathogenic".

Negative filters (remove):

	Gene	Not associated with a gene OR Gene starts with MUC, HLA, LINC
OR	VCF filter	QD < 2.0 ReadPosRankSum < -8. FS > 60.0 MQRankSum < -12.5
OR	Protein effect	Synonymous with no splicing effect OR Intron, UTR, upstream, downstream (not exonic) 　-　UNLESS -50 to +10 with splicing effect 　　–　OR disrupting 5' UTR [256]
OR	Frequency	Esp, 1000 genomes or gnomAD >=0.01

B Inheritance specific filters

Trio (assumption is two unaffected parents)

Mechanism	Child	Parents
Homozygous	Homozygous	Both heterozygous
Compound heterozygous	Compound heterozygous	Both heterozygous
De novo	Heterozygous	Both reference
Parental mosaicsm	Heterozygous	One parent mosaic
XLR	Hemizygous	Mother heterozygous
UPD	Homozygous	One parent only heterozygous
Imprinted	Variant in imprinted gene list	

Couple (as trio, but without child)

Mechanism	Parents
Homozygous	Both heterozygous
Compound heterozygous	Both heterozygous
De novo	Both reference
Parental mosaicsm	One parent mosaic
XLR	Mother heterozygous
UPD	One parent only heterozygous
Imprinted	

Two affected siblings (relatives)

Mechanism	Sibling1	Sibling2
Shared homozygous	Homozygous	Homozygous
Shared heterozygous	Heterozygous	Heterozygous

Table 2.6: Variant call format (VCF) file filtering based on inheritance

(**A**) Global filters for all exome data. (**B**) Inheritance specific filters including trios, couples and siblings. For families with autosomal recessive inheritance and autozygosity mapping data: regions of

homozygosity should be calculated and ranked, with each variant checked for location within these regions. Abbreviations: XLR, X-linked recessive; UPD, uniparental disomy; UTR, untranslated region.

Variant prioritisation

Variants within the WES filtered Excel files were manually reviewed and prioritised according to criteria listed in Table 2.7. Literature reviews were carried out in the assessment of the pathogenicity of genetic variants. Pubmed (https://pubmed.ncbi.nlm.nih.gov/), OMIM (https://www.omim.org/) and Human Gene Mutation Database (HGMD) Professional (https://digitalinsights.qiagen.com/product-login/) were used to assess available literature for specific genes and variants identified.

Parameter	Criteria for prioritisation
Call quality	VCF filter = pass
	Read depth ≥4
	Mapping quality ≥50
	Check variant authenticity through visualisation of aligned reads using the Integrative Genome Viewer (IGV)
Frequency in population databases	Mean allele frequency (MAF) <0.01 in:
	• GnomAD v2.1.1 and v3.1.2 (all populations)
	• 1000 Genomes Project
	• Internal control datasets for Amish and Pakistani individuals
	• Collaborative datasets (e.g. Amish Variant Server)
	Absence of homozygous individuals
Variant consistent with mode of inheritance in family pedigree	Autosomal recessive: homozygous or compound heterozygous
	Autosomal dominant: heterozygous
	X-linked recessive: hemizygous in males
Pathogenicity prediction tools (*in silico)*	Missense variants (non-synonymous):
	• PROVEAN score <-2.5
	• SIFT score <0.05
	• PolyPhen-2 score >0.85
	• REVEL score >0.7
	Splicing variants:
	• Splice AI (Broad Institute)
	• Human Splice Finder (HSF)
	• MaxEntScan
Predicted impact on protein or splicing	Protein-coding genes
	Protein truncating variants e.g. nonsense, frameshift
	Presumed loss of function variants e.g. small deletions, CNV's, structural
	Non-synonymous missense variants
	Splice variants within ±6 bases of the intron-exon junction
	Synonymous variants with an effect on splice site
	Variants disrupting a 5'UTR
Previously reported variants	DM (disease causing mutation) or DM? annotation in HGMDPro
	Pathogenic or likely pathogenic annotation in ClinVar
	Pathogenic or likely pathogenic annotation in Decipher
Variants in genes with a disease association	Genes listed as OMIM morbid genes with similarity between the associated disease and clinical phenotype of the family

Table 2.7: Variant prioritisation criteria

Abbreviations: VCF; variant call file, CNV; copy number variant, UTR; untranslated region, DM; disease causing mutation.

2.5 Generation of gene constructs

Gateway cloning technology (Invitrogen) is a rapid and highly efficient way to move DNA sequences into multiple vector systems. This is an *in vitro* version of the integration and excision recombination reactions that occur when lambda phage infects bacteria through attP and attB sites. The phage is flanked by two new sites attL and attR that can recombine leading to excision of the phage to regenerate attP and attB sites (Figure 2.2). Modifying these sites (e.g. attB1, attB2) allows for directional cloning.

Mutant constructs were created by Gateway Technology cloning using the pDONR™201 (Invitrogen) Gateway-adapted vector to generate an attL flanked entry clone containing the gene of interest by performing a BP recombination reaction with an attB PCR product. An expression clone was then created using an LR recombination reaction between the entry clone and destination vector. All work with live bacteria was carried out using a category 2 biological safety cabinet under sterile conditions with autoclaved sterile media. Assistance with generation of gene constructs was provided by Hashem Almousa (Concordia University) who created the entry clones.
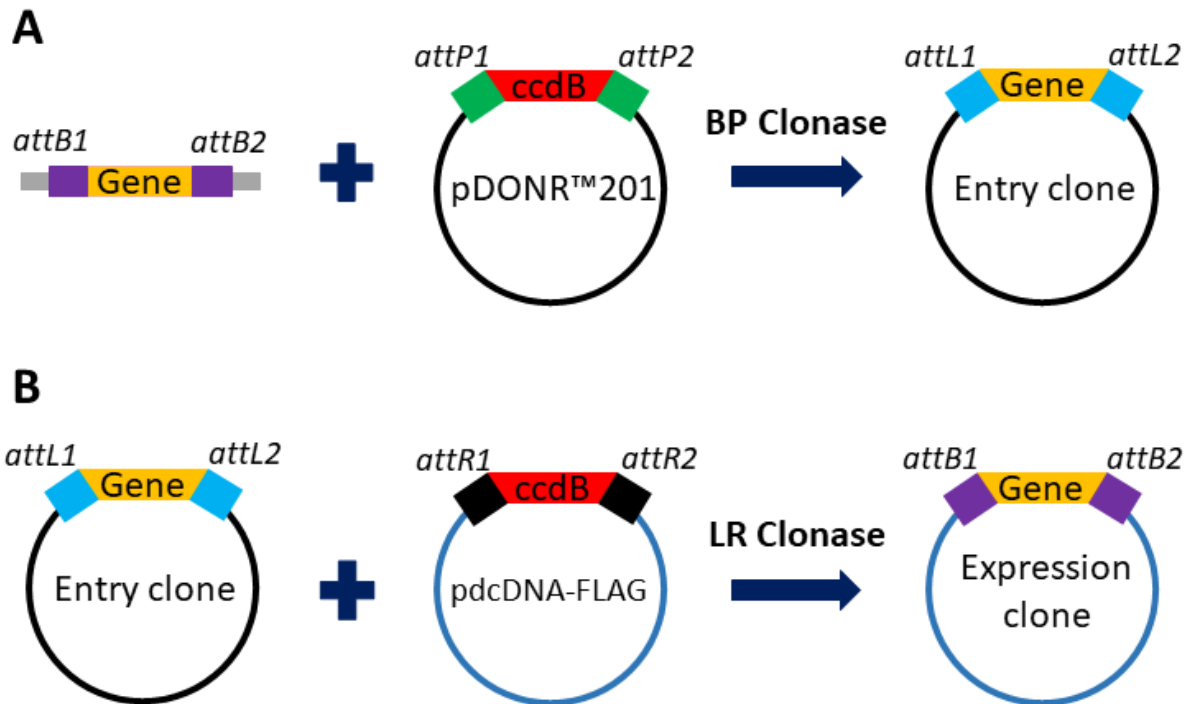
Figure 2.2: The gateway cloning system

(**A**) The BP reaction between attB sites flanking the insert and attP sites of the donor vector generates an attL-flanked entry clone. (**B**) The LR reaction between attL sites of the entry clone and attR sites of the destination vector generates an expression clone including the gene of interest. Adapted from Invitrogen Gateway Technology user guide 2003.

2.5.1 Generation of attB-flanked PCR products

attB-flanked PCR products containing the mutant constructs were created by Hashem Almousa (Concordia University) by site directed mutagenesis using oligonucleotides designed to contain a specific mutation in a gene. Table 2.8 lists the reaction components for the site directed mutagenesis PCR reaction.

Reagent (concentration)	Volume
Mutagenesis buffer (10x)	5µl
Plasmid DNA	2µl (50ng)
Forward primer	5µl (125ng)
Reverse primer	5µl (125ng)
dNTP mix (2mM)	2.5µl
ddH$_2$O	29.5µl
PfuUltra HF DNA polymerase	1µl (1 unit)
Total volume	50µl

Table 2.8: Reaction components for site directed mutagenesis PCR reaction

The reagents were added to a 0.2ml reaction tube, with 1µl PfuUltra HF DNA polymerase added last and placed in a thermocycler for the following protocol listed in Table 2.9:

Step	Temperature	Duration	Cycles
Incubation	37°C	60 mins	1
Initial denaturation	95°C	2 mins	1
Denature	95°C	20 sec	
Anneal	60°C	20 sec	18
Extension	72°C	4 min 8 sec	
Cool	4°C	Hold	1

Table 2.9: PCR protocol for site directed mutagenesis in pDONR™201

2.5.2 Entry clone creation

The pDONR™201 plasmid backbone vector contains attP1 and attP2 sites, a ccdB gene and a Kanamycin resistance gene (Appendix 1, Figure A1.1). Entry clones were generated using the following methods. The pDONR™201 vector was first resuspended in 40µl of sterile water (150ng/µl). pEXP7-tet consists of a linear sequence with attB sites flanking a tetracycline resistance gene and promotor and was used as a positive control. The attB PCR product and pDONR™201 vector were incubated with BP Clonase II in a 10µl reaction. The following were mixed in a 1.5ml microcentrifuge tube: 1µl of pDONR™201 vector, 2µl of attB PCR product and 5µl of TE buffer. A positive control was created using 2µl of pEXP7-tet instead of attB PCR product. A negative control was created by the addition of another 2µl of TE buffer. BP Clonase II enzyme was thawed and vortexed, then 2µl added to each sample and the positive control (not the negative control) followed by vortexing twice briefly. The mixture was incubated at 25°C overnight and 1µl of proteinase K was added before further incubation at 37°C for 10 minutes to terminate the reaction. Entry clones were then transformed into competent DH5α *Escherichia Coli.*

2.5.3 Bacterial transformation

Competent DH5α *E.Coli* were removed from the -80°C freezer and thawed on ice. 2.5µl of the BP recombination reaction was added to 50µl of DH5α *E.Coli* in a 1.5ml microcentrifuge tube and mixed gently before incubating on ice for 30 minutes. The cells were heat-shocked for 30 seconds in a 42°C water bath, before immediately transferring onto ice for 2 minutes. 1000µl of Luria Bertani (LB) media was then added and the tube capped and shaken for 1 hour at 37°C. The tube was then centrifuged at 2000rpm for 2 minutes so the pellet is just visible, then 850µl LB media removed from the top to leave ~200µl in which the cells were then resuspended. This 200µl from each transformation was then spread onto LB-Kanamycin (50µg/ml) agar plates using Pasteur pipettes. The positive pEXP7-tet control was spread onto an LB-tetracycline (20µg/ml) agar plate. The plates were allowed to dry then incubated at 37°C overnight to allow bacterial colonies to grow. A single colony was then transferred on a pipette tip to 3ml of LB media with Kanamycin in a 10ml centrifuge tube and incubated for 16 hours overnight before purifying the plasmid DNA.

2.5.4 Purification of plasmid DNA

Plasmid DNA was purified using the EZ-10 Spin Column Plasmid DNA MiniPrep kit (BioBasic). 1.5ml of LB medium cultured overnight (the remaining LB medium was saved as a stock at 4°C) was added to a 1.5ml microcentrifuge tube and centrifuged at 12,000rpm for 2 minutes to form a pellet. The supernatant was discarded, then 100µl of Solution I added to the pellet and mixed before incubating for 1 minute. 200µl of Solution II was then added and the tube inverted 6 times to mix before incubating for 1 minute. Then 350µl of Solution III was added, mixed and incubated for 1 minute, with the visible formation of white material and reduction in viscosity. The sample was then centrifuged at 12,000rpm for 5 minutes before transferring to the EZ-10 column and centrifuging for a further 2 minutes at 10,000rpm. Flow through was discarded and then 750µl of Wash Solution added to the column before centrifuging at 10,000rpm for 2 minutes. This wash step was repeated for a total of 2 washes. Flow through was discarded and the column centrifuged at 10,000rpm for 1 minute to remove residual wash solution. The column was then transferred to a clean, autoclaved 1.5ml microcentrifuge tube and 50µl of Elution Buffer added into the central part of the column membrane and incubated for 2 minutes. The tube and column were then centrifuged at 10,000rpm for 2 minutes before discarding the column and saving the eluted purified DNA, which was stored at -20°C.

Plasmid DNA concentration and purity were then measured using a NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific) as previously described. The presence of mutant gene constructs within the plasmid DNA was sequence verified by dideoxy sequencing.

2.5.5 Expression clone creation

Mutant constructs were FLAG-tagged for later use in cell studies, using an ampicillin resistant FLAG plasmid (pdcDNA-FLAG (LMBP 4704) Appendix 1, Figure A1.2). Expression clones were generated by recombination between an entry clone (containing the gene construct of interest) flanked by attL sites and the destination vector (FLAG plasmid) containing attR sites.

FLAG plasmid destination vector and purified plasmid DNA from mutant construct entry clones were incubated with LR Clonase II enzyme mix in a 5µl reaction. The following were mixed in a 1.5ml microcentrifuge tube: 1µl of FLAG plasmid, 1µl of entry

clone purified plasmid DNA and 1µl of ddH$_2$O. A negative control was created by the addition of another 2µl of TE buffer (for two entry clones). Then the LR Clonase II enzyme was thawed and vortexed, then 2µl added to the mixture (except the negative control) followed by vortexing and centrifuged briefly. The mixture was incubated at 25°C overnight and 1µl of proteinase K was then added before further incubation at 37°C for 10 minutes to terminate the reaction.

The resulting expression clones were then transformed into DH5α *E.Coli* and grown on LB-ampicillin (100µg/ml) agar plates to confirm the presence of the FLAG tagged plasmid, followed by colony selection and purification of plasmid DNA, according to methods previously described. The presence of the gene construct within the plasmid was then confirmed by restriction digest.

2.5.6 Restriction Digest

To confirm the presence of a specific gene sequence within a plasmid, restriction digestion was carried out. Restriction sites present in plasmid gene constructs that could be targeted with restriction enzymes were identified by pasting the DNA sequence of the gene into the NEBcutter V2.0 website (http://nc2.neb.com/NEBcutter2/) [257]. Before a restriction enzyme was selected the expected product sizes resulting from cleavage were identified. The resulting fragments were then visualised using agarose gel electrophoresis to confirm the presence or absence of the gene based on their size.

In a 1.5ml microcentrifuge tube 2µl of plasmid DNA was mixed with 6µl of dH$_2$O, 1µl of restriction enzyme buffer and 1µl of restriction enzyme. A negative control was created by the addition of 8µl dH$_2$O to 2µl plasmid DNA. The samples were incubated at 37°C for >1 hour. 2µl of 6x loading dye was added to each sample before loading onto a 1% agarose gel in 1x Tris borate EDTA (TBE) buffer containing 4µl of Redsafe™ Nucleic Acid Staining Solution. 4µl of a molecular weight marker (GeneRuler™ 1kb Plus DNA Ladder) was directly loaded into one well to assess the size of the products. The samples were electrophoresed through the gel at a constant voltage of 100V for 30 minutes before being visualised on a UV transilluminator and images recorded.

2.5.7 Ligation

Ligation of vector and insert DNA was performed using the Quick Ligation Kit and protocol (New England Biolabs). A 10μl reaction was prepared in a 1.5ml microcentrifuge tube on ice by adding 5μl of quick ligase reaction buffer, 25ng of vector DNA, 18.75ng of insert DNA, 0.5μl of quick ligase and nuclease free water added to 10μl final volume. The mixture was pipetted up and down then briefly centrifuged and incubated at room temperature for 5 minutes. The samples were chilled on ice before transformation into DH5α *E.Coli* or stored at -20°C.

2.5.8 Preparing samples for storage

5μl of resuspended stock LB medium containing DH5α with FLAG-tagged pDONR201 plasmid was cultured for freezing using methods previously described and incubated overnight. 1.5ml of LB medium was transferred to a microcentrifuge tube, then centrifuged at 2500rpm for 2 minutes to form a pellet. LB medium was poured off and 1μl of freezing medium consisting of 20% glycerol in LB medium was added before transferring to a screw top tube for storage at -80°C.

2.6 Cell culture techniques

HAP1 knockout cell lines used in subsequent parts of this study were created by Hashem Almousa (Concordia University) using CRISPR/Cas9 gene editing (Appendix 2, A2.1.1). Human lymphoblastoid cell lines were Epstein Barr virus (EBV) transformed by Public Health England, with lymphoblastoid cell studies performed by Hashem Almousa, Concordia University (sections 3.4.2.3 and 3.4.2.4).

2.6.1 Cell line culture

All cell line culture was performed in a tissue culture hood in a designated tissue culture room. All surfaces, consumables and reagent bottles were cleaned with 70% ethanol prior to use. All cells were maintained within appropriate cell culture medium in 37°C incubators. HeLa cells were maintained in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS). HAP1 cells were grown in Iscove's modification of Dulbecco's modified eagle medium (IMDM) with 10% FBS, 100U/ml penicillin and 100U/ml streptomycin. Cells were passaged once they became confluent by removal of cell culture medium and adding 3ml of trypsin (0.1%) and

incubating for 1-2 minutes to separate cells from the dish and then removing all trypsin and resuspending a small fraction of the cells in new tissue culture medium.

2.6.2 DNA transfection

DNA plasmids were transfected into cultured cells using jetPRIME® *in vitro* DNA and siRNA transfection reagent. Cells at 60-70% confluence in 6 well plates or 10cm dishes were used, with reagent quantities listed in Table 2.10.

Culture vessel	Volume of jetPRIME® buffer (µl)	Amount of DNA (µg)	Volume of jetPRIME® reagent (µl)	Volume of growth medium (ml)
6 well plate	200	2	4	2
10cm dish	500	10	20	10

Table 2.10: Quantities of DNA transfection reagents

For HeLa cells the amount of DNA was reduced to 0.5x as per protocol recommendations. A 1:2 ratio of DNA:jetPRIME® was used. First jetPRIME® buffer was added to sterile 1.5ml microcentrifuge tubes, followed by the correct quantity of DNA and mixed by vortexing. The appropriate volume of jetPRIME reagent was then added, vortexed for 1 second and briefly centrifuged before incubating at room temperature for 10 minutes. Fresh culture medium was added to the cells before the transfection mix was added dropwise onto the cells to distribute evenly and the plates gently rocked from side to side to mix before incubating at 37°C overnight. Cells were checked and culture medium changed the next morning, with lysate collected from cells after 24-48 hours.

2.6.3 Neomycin selection curve

A neomycin selection curve was carried out for HAP1 parental and knockout cell lines to identify the minimum concentration of neomycin needed to kill all cells over 7-10 days. Neomycin/G418 at 735µg/mg potency was used to create a 100x stock solution (100mg/ml) by dilution of 136.1mg in 1ml ddH$_2$O. This was then filtered using a 10ml lock syringe and a 0.2µm filter in the tissue culture hood before transferring to a sterile

1.5ml microcentrifuge tube and stored at -20°C. The stock 100x G418 was then diluted to 10x (10mg/ml) for use within the neomycin curve. 12 well plates were used, with HAP1 parental or knockout cells at ~30% confluence to start the antibiotic curve. IMDM culture medium and 10x G418 were added to each 1ml well in the volumes shown in Figure 2.3 and incubated at 37°C. The antibiotic containing medium was changed every 2-3 days and cell survival monitored daily for 10 days. For HAP1 parental and knockout cell lines all cells were dead at a concentration of 500µg/ml G418 and above after 10 days.

Neomycin/G418 selection (at 500µg/ml) was carried out for transfected HAP1 parental and knockout cells for 10 days to select for colonies of transfected cells. Surviving single colonies were trypsinised and transferred to separate 12 well dishes with IMDM culture medium containing 500µg/ml G418.

Figure 2.3: Neomycin survival curve

Neomycin survival curve showing volumes of 10x Neomycin (G418) and Iscove's modification of Dulbecco's modified eagle medium (IMDM) added to each of 12 wells.

2.6.4 Freezing cell lines

Cell culture medium was prepared with 5% DMSO and mixed thoroughly. Cells at ~80% confluency were prepared for freezing. Culture medium was removed and cells washed with 5ml of phosphate buffer solution (PBS) before adding trypsin for 2-3 minutes to separate cells. 10ml of fresh culture medium was added and the tube then centrifuged at 13,000rpm for 5 minutes, the medium was then discarded and the pellet resuspended in culture medium with 5% DMSO (1ml per sample to be stored). 1ml was then transferred to each cryotube and placed in a quick freeze container at -20°C before transferring to a -80°C freezer for 2 weeks before long-term storage in liquid nitrogen.

2.6.5 Cell lysis preparation for SDS-PAGE and Western blotting

To analyse the expression levels of proteins within cultured cells they were seeded in 10cm diameter dishes and lysed when 80% confluence was reached. Lysis buffer is a mixture of detergent and protease inhibitors, used with mechanical forces to breakdown the cell wall/membrane but not proteins within cells to generate a lysate. Lysis buffer was prepared as detailed in Table 2.2. Cell culture medium was removed from culture plates and cells washed twice with 5ml of 1x PBS chilled to 4°C, to remove any remaining FBS and dead cells. 350µl of lysis buffer was then added to the cells by dropping over the whole plate and leaving for 2-3 minutes to incubate and allow cells to lyse. A plastic cell lifter/scraper was then used to ensure all cells and debris were removed from the dish and transferred to a sterile 1.5ml microcentrifuge tube on ice at 4°C for 10-15 minutes before briefly vortexing twice. The tubes were then centrifuged at 13,200rpm for 20 minutes at 4°C to form a pellet of cell debris, then the supernatant was transferred to a sterile microcentrifuge tube and stored at -20°C.

2.7 Protein studies

Proteins within a sample were detected and analysed using sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) to separate proteins based on size followed by protein identification using antibodies directed towards a specific target protein.

2.7.1 Protein quantification - Bradford protein assay

The concentration of protein in cell lysates was quantified prior to SDS-PAGE using the Bradford Protein Assay (BioRad) on a spectrophotometer. Bovine serum albumin (BSA, 1mg/ml) was used to create a standard curve using 0, 2, 5 and 10µg of BSA in a final volume of 100µl in 1.5ml microcentrifuge tubes. Protein samples were made from 1µl of sample added to 99µl of dH$_2$O. A 1x BioRad protein assay reagent solution was made from a 1:4 dilution of 5x BioRad protein assay reagent concentrate and deionised water. The 100µl BSA and protein samples were then transferred to plastic cuvettes and mixed with 1ml of the 1x BioRad protein assay reagent solution before incubation at room temperature for 5 minutes. The absorbance at 595nm wavelength was then determined using a spectrophotometer (NanoDrop™ One/OneC) to calculate the protein concentration in µg/ml. If out of range, the sample was diluted by 1:10 then repeated.

Cell lysates were prepared for SDS-PAGE by making 15µl samples containing 30-50µg of protein. In 1.5ml heat resistant microcentrifuge tubes lysates were added (for 50µg this was calculated by 50 / protein concentration (µg/ml)) and made up to 15µl with ddH$_2$O. Sample buffer was prepared by adding 5% β-mercaptoethanol to 4x sample buffer containing loading dye (see Table 2.2 for preparation of 4x sample buffer) stored at -20°C. For example 25µl β-mercaptoethanol added to 475µl of 4x sample buffer. 5µl of sample buffer was added to each lysate and the tubes heated to 100°C on a heat block for 3 minutes to denature the samples. The samples were then allowed to cool and briefly centrifuged before loading on the gel, or were stored at -20°C.

2.7.2 SDS Polyacrylamide gel electrophoresis (PAGE)

SDS-PAGE was used to separate proteins based on their size. The percentage of acrylamide within the gel varied depending on the size of the proteins to separate (Table 2.11).

Protein size (kDa)	Acrylamide gel percentage
>80	8%
46-80	10%
25-46	12%
<25	15%

Table 2.11: Percentage of acrylamide gel used based on size of protein

Mini-Protean TGX™ precast polyacrylamide 8-16% gels were used or 8% gels prepared by Djennan Saint-Dic (Concordia University) using the following methods. To make 2 gels (20ml), first the resolving gel was prepared by mixing in order 9.47ml ddH$_2$O, 5ml of 4x resolving buffer (see Table 2.2 for composition), 5.33ml of 30% acrylamide-bisacrylamide, 200µl of 10% ammonium per sulphate (APS) and 12µl of TEMED (N,N,NN-tetramethylethylenediamine), with 7-8ml poured into the gel casting apparatus. 700µl of isopropanol/ethanol was used to achieve a straight edge and the gel was left to polymerise for 30 minutes. The isopropanol was then removed before adding the stacking gel, which was prepared by mixing in order 1.79ml ddH$_2$O, 750µl of 4x stacking buffer (see Table 2.2 for composition), 400µl of 30% acrylamide, 30µl of 10% APS and 3µl of TEMED. 3-4ml of stacking gel was then added to the top of the gel, a comb inserted to create loading wells and was left to polymerise for 30 minutes. The electrophoresis frame was assembled and either two gels or one gel and a plastic insert were placed in the tank.

First, 1L of 1x running buffer was prepared by adding 100ml of 10x running buffer (see Table 2.2 for composition) to 895ml of dH$_2$O, then adding 5ml of 20% SDS for 0.1% final concentration. Running buffer was added to the middle of the well and filled to the top, bubbles were removed from the top and the tank was checked for leaks. 4µl of a molecular weight marker (colour protein standard) was added to the left hand lane and 20µl lysate samples added to each well. The cover was added and the samples fractionated at 100V for 90 minutes.

2.7.3 Transfer to membrane

After separation of proteins they were transferred to a nitrocellulose membrane. Transfer buffer (1x) was prepared by mixing 100ml of 10x transfer buffer (see Table

2.2 for composition), 700dH$_2$O and 200ml of methanol in a fume hood. The gel was equilibrated with transfer buffer for 10 minutes, then 4 pieces of chromatography grade blotting paper and one piece of nitrocellulose membrane cut to size were soaked in transfer buffer along with two sponges. One sponge with two pieces of blotting paper and the nitrocellulose membrane were rolled out to remove bubbles, then the gel (with stacking gel removed) placed on top with two pieces of blotting paper and then the second sponge, before rolling again to remove bubbles. The sandwich was clipped shut and placed in the tank with an ice block in the front of the chamber and transfer buffer added to the top of the tank before transferring at 100V for 1 hour.

2.7.4 Western blotting

Following transfer of the proteins to the nitrocellulose membrane, the membrane was placed in blocking solution to prevent non-specific antibody binding to the membrane. First 1x PBST was made by adding 50ml of 10x PBS to 450ml of ddH$_2$O, then adding 250µl of Tween 20 detergent. This was added to milk powder to make a 5% blocking solution (for example 5g in 100ml). The edges of the membrane were trimmed then placed in the 5% blocking solution and gently shaken for 1 hour.

The membrane was removed from the blocking solution, then transferred to a smaller container with 3ml of PBST and primary antibody added at the appropriate concentration (Table 2.12) and incubated on a shaker for 1 hour at room temperature (or overnight in a cold room). The primary antibody was poured off then the membrane was washed three times with 1x PBST for 5 minutes on the shaker to remove excess antibody. Then 3ml of PBST was added, with the secondary antibody added at the appropriate concentration (Table 2.12) and incubated at room temperature for 30-60 minutes on a shaker. The secondary antibody was poured off then washed three times with 1x PBST for 5 minutes.

Primary antibodies		
Antibody	**Type**	**Dilution**
Anti-TRAPPC10 (Santa Cruz Biotechnology, sc101259)	Mouse monoclonal IgG	1 in 200
Anti-FLAG (Sigma-Aldrich, F3165)	Mouse monoclonal	1 in 5000
Anti-alpha tubulin (Sigma-Aldrich, T6199)	Mouse monoclonal	1 in 8000
Secondary antibodies		
Anti-mouse IgG	Horseradish peroxidase conjugated Host: Rabbit	1 in 5000 1 in 10,000

Table 2.12: Primary and secondary antibodies used in Western blotting

Protein bands on the nitrocellulose membrane were detected using chemiluminescence. Amersham ECL™ Start Western Blotting Detection Reagent was used. A working solution was created immediately prior to use by mixing equal amounts (1ml of each) of detection solution A (luminol) and solution B (horseradish peroxidase). Any excess PBST wash was drained and enough working solution added to cover the membrane and incubated for 90 seconds at room temperature before draining off any excess with tissue paper. The membrane was then placed between two sheets of plastic and bubbles removed before visualising with an Amersham 600 CCD camera (GE Healthcare Life Sciences). Exposure was started at 30 seconds then increased to 1, 3 and 10 minutes depending on the resulting image.

3 BIALLELIC VARIANTS IN *TRAPPC10* CAUSE A MICROCEPHALIC TRAPPOPATHY DISORDER IN HUMANS AND MICE

3.1 Summary

The highly evolutionarily conserved transport protein particle (TRAPP) complexes (TRAPP II and III) perform fundamental roles in subcellular trafficking pathways. Here we identified biallelic variants in *TRAPPC10*, a component of the TRAPP II complex, in individuals with a severe microcephalic neurodevelopmental disorder. Molecular studies revealed a weakened interaction between mutant TRAPPC10 and its putative adaptor protein TRAPPC2L. Studies of patient lymphoblastoid cells revealed an absence of TRAPPC10 alongside a concomitant absence of TRAPPC9, another key TRAPP II complex component associated with a clinically overlapping neurodevelopmental disorder. The TRAPPC9/10 reduction phenotype was recapitulated in *TRAPPC10*$^{-/-}$ knockout cells, which also displayed a membrane trafficking defect. Notably, both the reduction in TRAPPC9 levels and the trafficking defect in these cells could be rescued by wild type but not mutant TRAPPC10 gene constructs. Moreover, studies of *Trappc10*$^{-/-}$ knockout mice revealed neuroanatomical brain defects and microcephaly, paralleling findings seen in the human condition as well as in a *Trappc9*$^{-/-}$ mouse model. Together these studies confirm autosomal recessive *TRAPPC10* variants as a cause of human disease and define TRAPP-mediated pathomolecular outcomes of importance to TRAPPC9 and TRAPPC10 mediated neurodevelopmental disorders in humans and mice.

Our defining *TRAPPC10* variants as a cause of human disease prompted a wider evaluation of other TRAPP complex molecules as candidate new causes of human disease. This resulted in the identification of 11 further families affected with neurodevelopmental disorders with features overlapping those of TRAPPopathies, associated with candidate biallelic variants in *TRAPPC1*, *TRAPPC3*, *TRAPPC5*, *TRAPPC8* and *TRAPPC14* genes. Together this consolidates TRAPPopathy disorders for which only single families have been previously described (*TRAPPC3*, *TRAPPC14*), and delineates potential new candidate TRAPPopathy disorder genes (*TRAPPC1*, *TRAPPC5*, *TRAPPC8*) for further investigation.

3.2 Introduction

The Transport Protein Particles (TRAPPs) comprise highly evolutionarily conserved multiprotein complexes, originally identified in yeast, which play important roles in secretory and endocytic subcellular trafficking pathways [258, 259]. Three yeast TRAPP complexes have been described TRAPP I, II and III [258, 260]. Humans possess only two complexes (TRAPP II and III), both sharing the same core proteins (TRAPPC1, TRAPPC2, TRAPPC2L, TRAPPC3, TRAPPC3L, TRAPPC4, TRAPPC5 and TRAPPC6) that comprise the yeast TRAPP I complex, alongside additional complex-specific components, some not identified in yeast (Figure 3.1) [261]. In addition to the core proteins TRAPP II contains two (TRAPPC9 and TRAPPC10) and TRAPP III contains four (TRAPPC8, TRAPPC11, TRAPPC12 and TRAPPC13) unique complex-specific subunits [105]. An additional TRAPP II complex-associated protein has recently been described (TRAPPC14/C7orf43) [262]. However, TRAPPC14 does not appear to be required for TRAPP II complex-associated guanine nucleotide exchange factor activity [263, 264] and lack of the molecule does not affect TRAPP II complex assembly or stability [262]. Thus, further studies are required to clarify the role of TRAPPC14 in TRAPP II.

Both TRAPP II and III complexes have been implicated in diverse cellular roles including vesicle tethering processes involved in the early secretory pathway, endocytosis and autophagy [260, 265-268], with a role in centrosomal functions and ciliogenesis [262, 269, 270]. Several TRAPP subunit proteins have been identified to have additional molecular functions independent of the role of TRAPP II and III complexes, for example TRAPPC2 regulates the activity of a transcriptional repressor (MBP1) [271]. The molecular actions of the TRAPP II complex include guanine nucleotide exchange factor (GEF) activity towards Rab GTPases [266, 267], which act as regulatable switches to recruit effector molecules that control intracellular vesicular trafficking. TRAPP III components have also been identified to play a role in protein glycosylation and lipid linked oligosaccharide biosynthesis in humans [272, 273].
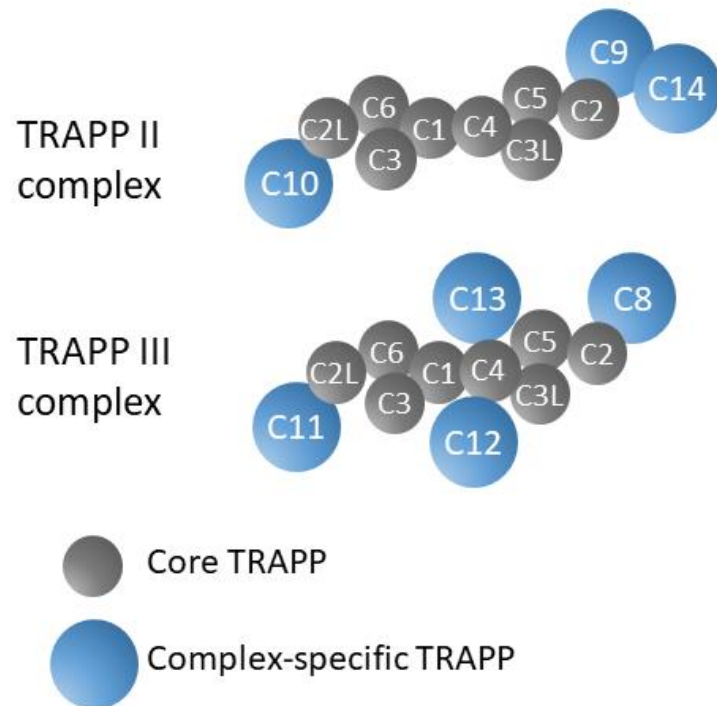
Figure 3.1: Schematic figure of the two human TRAPP complexes: TRAPP II and TRAPP III

Core subunits are found in both complexes and shown in grey, with complex-specific subunits coloured in blue. The arrangements of the subunits is based on known subunit organisation and interactions. Adapted with permission from Sacher *et al* [105] (licence granted by John Wiley and Sons, Appendix 7).

Within the TRAPP II complex, TRAPPC9 has been implicated in COP II vesicle trafficking [274] and NF-κB signalling pathway activation [275], with TRAPPC9 depletion associated with increased lipid droplet formation [267]. TRAPPC14 plays a role in Rabin 8 binding to the TRAPP II complex and its tethering of preciliary vesicles to the mother centriole during ciliogenesis [262]. No detailed studies of TRAPPC10 function have been conducted, although it has been shown to interact with TRAPPC2L and may have roles in GEF activity regulation [276].

Ten of the 16 identified TRAPP complex proteins have been associated with human inherited diseases, collectively termed "TRAPPopathies" [105]. This includes the well characterised TRAPPC2; the only subunit associated with skeletal dysplasia (X-linked spondyloepiphyseal dysplasia tarda (SEDT) OMIM #313400) [277]. The underlying pathomechanism of this condition involves the disruption of TRAPPC2-specific function in the export of collagen from the endoplasmic reticulum (ER) [278], and this distinct skeletal phenotype is thought to arise from a specific non TRAPP-complex function [105]. The remaining nine TRAPP complex proteins (TRAPPC2L, TRAPPC4, TRAPPC6A, TRAPPC6B, TRAPPC9, TRAPPC10, TRAPPC11, TRAPPC12 and TRAPPC14) have been associated with autosomal recessive neurodevelopmental disorders. While diverse, the clinical features of these disorders are overlapping and include neurodevelopmental delay, intellectual disability, developmental regression, primary and postnatal-onset microcephaly, epilepsy, hypotonia, craniofacial dysmorphism and structural brain abnormalities including thinning of the corpus callosum and reduced white matter volume (Table 3.1). A more detailed comparison is provided in Appendix 3, Table A3.1.

TRAPP subunit	Inheritance	OMIM number and clinical features	References
TRAPPC2	XLR	**Spondyloepiphyseal dysplasia tarda (OMIM #313400)** Short stature, vertebral flattening, kyphosis/scoliosis, small capital femoral epiphyses, mild epiphyseal irregularities	[279],[277] many see GeneReview [140]
TRAPPC2L	AR	**Encephalopathy, progressive, early-onset, with episodic rhabdomyolysis (OMIM #618331)** Developmental delay, absent speech, microcephaly, epilepsy, regression, rhabdomyolysis, raised CK, tetraplegia, dystonia, cerebral visual impairment	[276],[280]
TRAPPC3	AR	Developmental delay, intellectual disability, obesity, poor night vision, polydactyly, nystagmus, dextrocardia, complex cardiac disease	Single family described [281]
TRAPPC4	AR	**Neurodevelopmental disorder with epilepsy, spasticity, and brain atrophy (OMIM #618741)** Early-onset seizures, developmental delay, microcephaly, sensorineural deafness, spastic quadriparesis, progressive cortical and cerebellar atrophy, scoliosis, facial dysmorphism	[282], [283], [284], [285], [286]
TRAPPC6A	AR	Intellectual disability, speech delay, facial dysmorphism, polydactyly	Single family described [287]
TRAPPC6B	AR	**Neurodevelopmental disorder with microcephaly, epilepsy, and brain atrophy (OMIM #617862)** Microcephaly, intellectual disability, epilepsy, ataxic gait, cortical atrophy, autistic features, generalised weakness, thin corpus callosum	[288], [289], [290], [291]
TRAPPC9	AR	**Mental retardation, autosomal recessive 13 (OMIM #613192)** Microcephaly, intellectual disability, reduced white matter volume, thin corpus callosum, absent speech, hypotonia, dysmorphic features, hyperkinesia, epilepsy, raised BMI	[292], [293], [294], [295], [296], [297], [298], [299], [300], [301], [302], [303], [304], [305], [306]
TRAPPC10	AR	Microcephaly, intellectual disability, developmental delay, speech delay, hypotonia, gait abnormalities, epilepsy, autistic features, facial dysmorphism, short stature, strabismus, thin corpus callosum, reduced white matter structures	[307] *This study*
TRAPPC11	AR	**Muscular dystrophy, limb-girdle, autosomal recessive 18 (OMIM #615356)** Epilepsy, myopathy, microcephaly, developmental delay, ataxia, cerebral atrophy, raised CK, scoliosis, cataracts, hepatomegaly, steatohepatitis, hypotonia, gait abnormalities, short stature, dystonia, spasticity, generalised weakness, achalasia, alacrima	[308], [309], [310], [311], [273], [312], [313], [314], [315], [316], [272]
TRAPPC12	AR	**Encephalopathy, progressive, early-onset, with brain atrophy and spasticity (OMIM #617669)** Hypotonia, microcephaly, epilepsy, global developmental delay, brain atrophy, agenesis of the corpus callosum, pons hypoplasia, spasticity, myoclonic jerks, neurogenic bladder	[317], [318], [319]
TRAPPC14	AR	**?Microcephaly 25, primary, autosomal recessive (OMIM #618351)** Short stature, microcephaly, global developmental delay, impaired intellectual development, speech delay, ADHD, decreased white matter and thin corpus callosum	Single family described [320]

Table 3.1: Summary of the core clinical features associated with TRAPPopathy disorders, including TRAPPC10 defined here

Genes shaded in grey indicate candidate disease genes with only single families reported and limited functional evidence for pathogenicity of variants.

Abbreviations: TRAPP; transport protein particles. XLR; X-linked recessive, AR; autosomal recessive, BMI; body mass index, CK; creatine kinase.

Of the TRAPP II complex specific subunits, only TRAPPC9 has been robustly associated with human disease, an autosomal recessive neurodevelopmental disorder (OMIM #613192) characterised by postnatal-onset microcephaly with reduced white matter volume and corpus callosum thinning, intellectual disability, dysmorphic features, hypotonia, epilepsy, and raised BMI [292-295]. TRAPP II plays important roles in Golgi membrane trafficking. "Golgipathy" neurodevelopmental disorders including the TRAPPopathies comprise an expansive group of conditions associated with variants in Rab GTPases, conserved oligomeric Golgi (COG) complex, and coat protein (COP) complex proteins among others [7, 106] and are characterised by primary or postnatal-onset microcephaly, intellectual disability, seizures and white matter brain abnormalities. Numerous membrane trafficking disorders have been described [321], frequently characterised by neurodevelopmental features, including microcephaly (see Figure 1.3).

This chapter entails the description of comprehensive genetic, clinical, functional and mouse data to define biallelic *TRAPPC10* gene variants as a cause of a microcephalic neurodevelopmental disorder, providing important further insight into the TRAPPopathy family of disorders. Stemming from this discovery, a wider evaluation of previously unconfirmed TRAPP complex disease genes was undertaken. Appendix 3 summarises the results of an extensive literature review and compilation of data from the DDD project (Complementary analysis project (CAP) #303) the 100,000 genomes project (Genomics England) and international collaboration through GeneMatcher [244]. Together these datasets support the identification of three candidate new TRAPPopathy disease genes associated with variants in *TRAPPC1*, *TRAPPC5* and *TRAPPC8*.

3.3 Materials and Methods

Detailed descriptions of the genetic, clinical and functional methodology for these studies are provided in Chapter 2.

Genetic studies

Genomic DNA was extracted from blood/buccal samples (section 2.4.1) and RNA was extracted and converted to cDNA via RT-PCR (sections 2.4.2 and 2.4.5). Genome-

wide SNP microarray genotyping and WES analysis was performed (NextSeq500; Illumina, San Diego, CA, USA) and involved Agilent SureSelect Whole Exome v6 (Agilent Technologies, Santa Clara, CA), see sections 2.4.6 and 2.4.7. Primers were designed (Appendix 1, Table 1.1) and dideoxy sequencing validation of *TRAPPC10* variants was undertaken using standard techniques (section 2.4.2) to confirm variant segregation with the disease phenotype.

Cellular/molecular studies

Molecular studies were carried out in collaboration with Professor Michael Sacher (Concordia University, Montreal). The aim of these studies was to ascertain the effect of *TRAPPC10* variants on molecular interactions, TRAPP II complex function and membrane trafficking. Many of these studies were undertaken by myself during a six month placement at Concordia University, working closely with Hashem Almousa, Ana Maria Hincapie, Dr Miroslav Milev and Dr Djennan Saint-Dic, researchers from the Sacher laboratory. Where specific experiments were performed by members of the Sacher laboratory this is detailed in the figure legends. The methodology for the functional studies performed by Professor Sacher's research group are detailed in Appendix 2, section A2.1.1.

Standard molecular biological techniques were used to generate FLAG-tagged pDONR™201 constructs and *TRAPPC10* variants as detailed in section 2.5. Variant TRAPPC10 constructs were transfected into HeLa cells using jetPRIME (section 2.6) to confirm their detection using Western blotting (section 2.7).

Confirmation of HAP1 TRAPPC10$^{-/-}$ cell lines (created by Hashem Almousa, Concordia University) included designing primers (Appendix 1, Table 1.1), performing PCR for dideoxy sequencing, and cell lysis with Western blotting, to confirm knockout of TRAPPC10 within HAP1 cells (section 2.7). I performed rescue experiments with transfection of a FLAG-tagged wildtype (WT) TRAPPC10 construct using jetPRIME (section 2.6), followed by neomycin selection (section 2.6.3).

Mouse studies

Mouse studies were performed by Dr Binnaz Yelcin's research group (Inserm, France) by Dr Stephan C Collins, Dr Christine Rowley, Dr Valerie E Vancollie and Dr Christopher J Lelliott. Detailed methodology is included in Appendix 2, section 2.1.2.

3.4 Results

3.4.1 Clinical and genetic findings

3.4.1.1 Biallelic *TRAPPC10* variants are associated with a microcephalic neurodevelopmental disorder

All affected and unaffected family members included in this study were recruited with written informed consent (see section 2.2.1) and phenotyping carried out according to methods detailed in section 2.2.2. Table 3.2 summarises the clinical features of affected individuals in this study. Family 1 comprises a large interlinking multi-nuclear Pakistani family, with eight individuals (4.1-18.1 years) affected by a severe microcephalic neurodevelopmental disorder (Figure 3.2A). All but one individual presented with microcephaly (>3 standard deviations (SDS) below mean), and all displayed short stature (-1.78 to -5.09 SDS). All eight were born at full term, prenatal histories were unremarkable, although antenatal ultrasound scans were not performed. The parents did not report that any affected infant's head was noticeably small (no weight or occipitofrontal circumference (OFC) measurements were recorded). The earliest OFC measurement documented was at 2.5 years (1-IV:4), and at age 4.1 years his microcephaly had not worsened.

FAMILY	FAMILY 1								FAMILY 2		FAMILY 3	FAMILY 4
Individual	IV:2	IV:3	IV:4	IV:7	IV:8	IV:10	IV:12	IV:13	IV:1	IV:2	II:1	II:1
Genotype	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(G1131Vfs*19)/p.(G1131Vfs*19)	p.(P929L)/p.(P929L)	p.(P929L)/p.(P929L)	p.(L502F)/p.(A673T)-p.(G899R)	p.(C1074*)/p.(C1074*)
Gender	M	M	M	F	F	F	M	M	M	M	F	F
Age at evaluation (years)	18.1	13.8	4.1	14.1	15.6	5.5	9.0	13.7	25	22	6	9
GROWTH												
Birth weight kg (SDS)	NK	NK	2.5 (-2.25)	NK	NK	NK	NK	NK	NK	NK	NK	NK
OFC cm (SDS)	55 (-1.32)	50 (-3.49)	47 (-3.59)	46 (-6.64)	49 (-4.56)	43 (-7.46)	43 (-7.05)	47 (-5.28)	50 (-4.24)	53 (-2.49)	NK	45 (-6.77)
Height cm (SDS)	164 (-1.87)	146 (-1.78)	94 (-2.23)	127 (-5.09)	142 (-3.41)	99 (-2.78)	119 (-2.48)	125 (-4.26)	150 (-3.92)	140 (-5.29)	NK	78 (-9.38)
Weight kg (SDS)	52 (-1.96)	40 (-0.99)	18 (0.6)	25 (-4.9)	45 (-1.39)	16 (-1.44)	20 (-2.77)	28 (-3.16)	45 (-3.9)	38 (-4.87)	NK	18.2 (-3.23)
BMI	19.3	18.7	20.4	15.5	22.3	16.3	14.1	17.9	20	19	NK	29.9
DEVELOPMENT												
Intellectual disability	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe	Severe
Global developmental delay	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	NK	✓
Speech impairment	✓ <10 words	✓ Non-verbal	✓ <10 words	✓ Non-verbal	✓ <10 words	✓ Non-verbal	✓ <10 words	✓ <10 words	✓	✓	NK	✓ Non-verbal
Walked (years)	4	4	Standing	3.5	2.5	4	4	3	NK	NK	NK	Non-ambulatory
Hearing	No concerns	Otitis media	No concerns	No concerns	No concerns	No concerns	No concerns	No concerns	No concerns	No concerns	NK	SNHL (*GJB2* variant)
Vision	No concerns	No concerns	No concerns	No concerns	No concerns	No concerns	Strabismus	No concerns	Strabismus	Strabismus	NK	No concerns
NEUROLOGY												
Childhood hypotonia	✓	✓	✓	✓	✓	✓	✓	✓	NK	NK	NK	✓
Seizures	✓	✓	-	✓	-	-	-	✓	-	-	NK	✓
Gait abnormalities	-	Waddling gait	-	-	-	-	Waddling gait	Waddling gait	NK	NK	NK	Non-ambulatory
Behavioural abnormalities	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table 3.2: A comparison of clinical features of affected individuals with biallelic variants in *TRAPPC10*

115

Height, weight and OFC Z-scores were calculated using a Microsoft Excel add-in to access growth references based on the LMS method [245] using UK1990 reference population [54]. (✓), indicates presence of a feature in an affected individual; (-), indicates absence of a feature in an affected individual. Grey shading indicates that the pathogenicity of the *TRAPPC10* variants were not confirmed in this family. Abbreviations: M, male; F, female; NK, not known; SDS, standard deviation score; OFC, occipitofrontal circumference; GDD, global developmental delay; SNHL, sensorineural hearing loss.
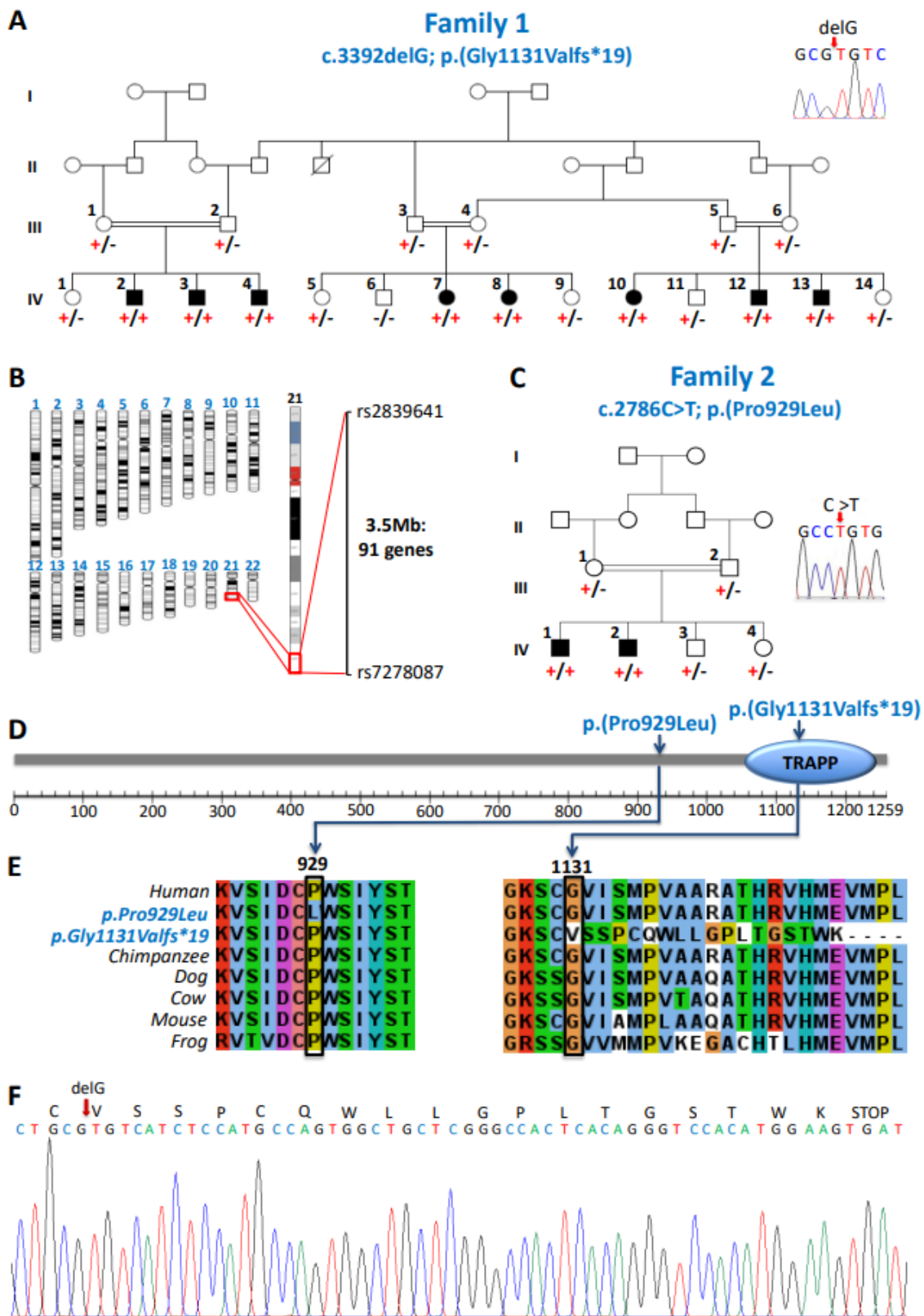
Figure 3.2: Biallelic *TRAPPC10* variants identified in two families with individuals affected by a microcephalic neurodevelopmental disorder

(**A**) Family 1: Extended pedigree comprising three interlinking families and a total of eight affected individuals all of whom are homozygous for the *TRAPPC10* c.3392delG; p.(Gly1131Valfs*19) variant ('+'), with co-segregation confirmed in other family members ('-' indicates the wild type allele). Inset: Dideoxy sequence chromatogram of an affected individual homozygous for the *TRAPPC10* c.3392delG; p.(Gly1131Valfs*19) variant. (**B**) Genome-wide SNP mapping in seven affected individuals from Family 1(IV:2-4, IV:7-8 and IV:12-13) identified a single (3.5Mb) region of shared homozygosity, containing 91 genes including *TRAPPC10*. (**C**) Family 2: Previously reported family (MR107) [27] showing co-segregation of the *TRAPPC10* c.2786C>T; p.(Pro929Leu) variant ('+') in a four generation pedigree ('-' indicates the wild type allele). Inset: Dideoxy sequence chromatogram of an affected individual homozygous for the *TRAPPC10* c.2786C>T; p.(Pro929Leu) variant. (**D**) Schematic diagram of the TRAPPC10 protein identifying *TRAPPC10* sequence variants and TRAPP domain. (**E**) Multi-species alignment showing conservation of the molecular region encompassing the p.Gly1131Valfs*19 and p.Pro929Leu variants. (**F**) Chromatogram of the *TRAPPC10* c.3392del rtPCR product, confirming the outcome of the variant on the RNA transcript in blood of an affected individual (1-IV:2), demonstrating that this frameshift variant results in a premature stop codon 19 codons downstream.

Affected individuals displayed mild craniofacial dysmorphism (Figure 3.3C-K). Although standardised intelligence quotient (IQ) testing was not possible due to cultural and language barriers, all affected individuals were assessed as severely intellectually impaired (DSM-V criteria). Developmental trajectories were similar in all, as was the degree of impairment across all developmental domains. Hypotonia in infancy was universal and walking was delayed to between 2.5-4 years. All affected individuals have speech impairment, five have only a few words and three are non-verbal. Behavioural abnormalities including autistic features and aggressive episodes were a consistent feature. A history of seizures from early infancy was reported in four individuals, one has ongoing generalised tonic-clonic seizures (1-IV:2) and the remaining three (1-IV:3, 1-IV:7, 1-IV:13) are seizure free on medication. Neuroimaging was only performed for one case (1-IV:2), due to the remote family location and as sedation/general anaesthesia was deemed necessary. This revealed microcephaly with corpus callosum thinning and no other abnormalities (Figure 3.3A,B). All had normal systemic examination findings and no evidence of gross skeletal disproportion (Figure 3.3C). It was only possible to obtain skeletal radiographs for a single individual (1-IV:12) aged 9 years, these were reviewed by Professor Christine Hall (Emeritus, Great Ormond Street Hospital) and found to be unremarkable but unfortunately spinal imaging was incomplete.
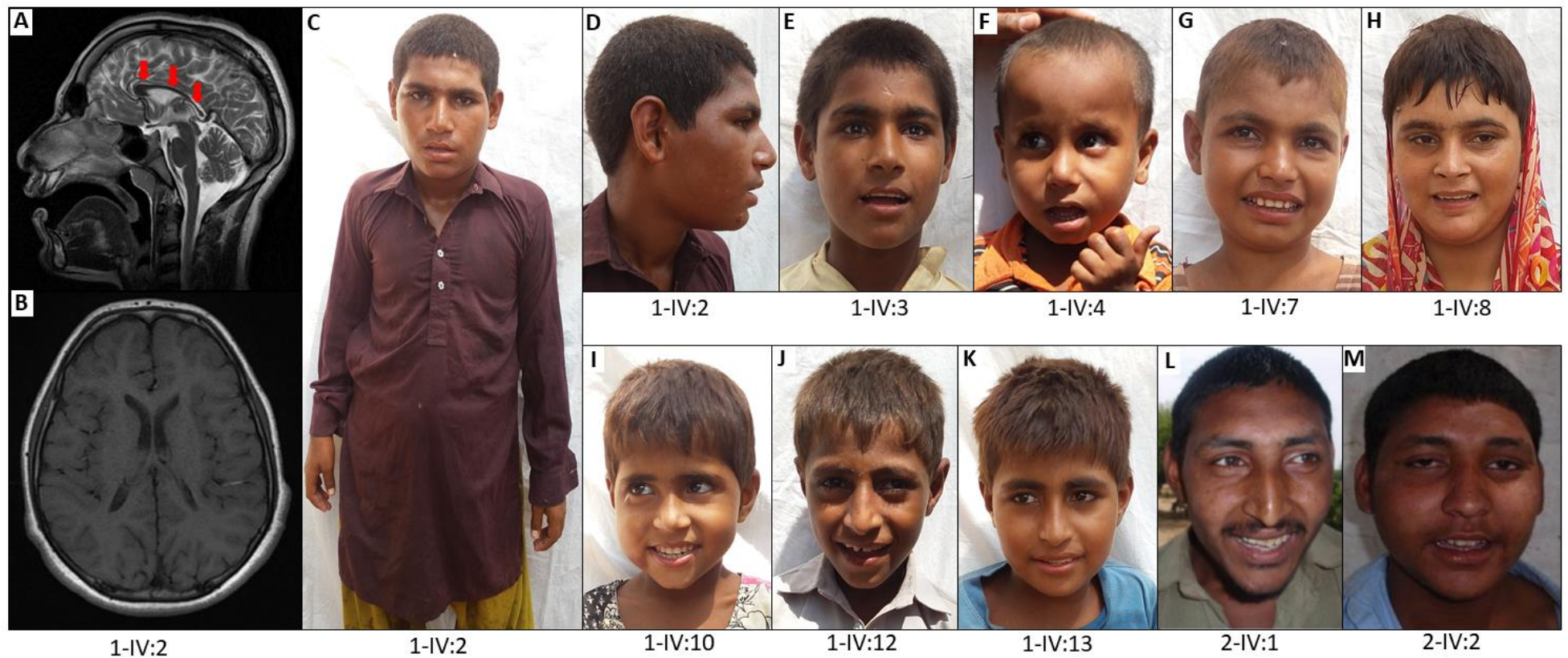
Figure 3.3: MRI neuroimaging and clinical photographs of individuals with biallelic *TRAPPC10* variants

(**A,B**) MRI neuroimaging of individual IV:2 from Family 1 demonstrates microcephaly and profound thinning of the corpus callosum (red arrows) with no other abnormalities. (**C**) Shows individual 1-IV:2 (Family 1) with no evidence of skeletal disproportion. (**D-K**) Show the facial features of all affected individuals from Family 1 displaying mild craniofacial dysmorphism comprising microcephaly, synophrys and upslanting palpebral fissures. (**L,M**) Show similar facial features of brothers from Family 2 [322].

Assuming homozygosity for a founder variant was responsible, we undertook SNP genotyping (1-IV:2-4, 1-IV:7-8, 1-IV:12-13) in parallel with exome sequencing (1-IV:2) to identify the cause of disease. SNP mapping identified a single notable (>1Mb) homozygous region common to all affected individuals (3.5Mb, rs2839641-rs7278087; chr21:g.43,179,611-46,678,912 [hg38]) as the likely disease locus (Figure 3.2B). Concomitant exome sequencing data analysis to identify candidate homozygous and compound heterozygous rare (<0.01 frequency) variants predicted to have a functional impact, identified a single candidate variant within this region which could not be excluded. This frameshift variant (chr21:g.44102823del; NM_003274.4: c.3392del; p.(Gly1131Valfs*19) [hg38]; Figure 3.2A) in the penultimate (22/23) exon of *TRAPPC10* located in the chromosome 21q22.12 region, is present within the C-terminal TRAPP domain (Figure 3.2D) and is predicted to result in nonsense mediated mRNA decay (NMD) of the mutant transcript (Figure 3.2E). The variant was confirmed in the lymphocyte RNA of an affected individual (Figure 3.2F), is not listed in publicly available genomic databases (including gnomAD v2.1.1/v3.1.2, Table 3.3) and co-segregates appropriately (Figure 3.2A).

This finding enabled us to revisit the genetic data and more comprehensively phenotype two siblings in an unrelated Pakistani family (Figure 3.2C) included in a larger study to identify novel candidate neurodevelopmental disease genes [307, 322]. Both male siblings (2-IV:1, 22 years and 2-IV:2, 25 years), were described to have "*severe intellectual disability, aggressive behaviour and poor speech*". A homozygous exon 18 *TRAPPC10* variant (chr21:g.44089849C>T; NM_003274.4:c.2786C>T; p.(Pro929Leu) [hg38]) (Figure 3.2C) was identified as a potential cause of disease [307]. This variant, which co-segregated appropriately, is absent from gnomAD v2.1.1/v3.1.1 (Table 3.3). Multiple species alignment confirms stringent conservation of Pro929 across vertebrates (Figure 3.2E) and multiple missense prediction tools suggest this variant is deleterious (Table 3.3). Additional clinical features noted and not previously reported, included microcephaly/borderline microcephaly (-2.49 SDS, -4.2 SDS), short stature (-3.86, – 5.29 SDS) and mild dysmorphism (Table 3.2, Figure 3.3L,M), all consistent with Family 1.

CASES	Variant	GRCh38 (hg38)	gnomAD v2.1.1 HET	gnomAD v3.1.2 HET	gnomAD HOM	gnomAD All AF	SIFT	Provean	Polyphen-2	REVEL	Splice AI
TRAPPC10: NM_003274.4											
Family 1 (HOM)	c.3392del; p.(Gly1131Valfs*19)	chr21:g.44102823del	-	-	-	-	NA				
Family 2 (HOM) Santos-Cortez et al	c.2786C>T; p.(Pro929Leu)	chr21:g.44089849C>T	-	-	-	-	0.005	-6.9	0.999	0.526	0.00
Family 3 (pat)	c.1504C>T; p.(Leu502Phe)	chr21:g. 44079598C>T	101	61	0	0.0004008	0.000	-1.29	0.997	0.421	0.06
Family 3 (mat)	c.2017G>A; p.(Ala673Thr)	chr21:g.44083081G>A	3	2	0	0.00001316	0.323	-0.61	0.016	0.045	0.01
Family 3 (mat)	c.2695G>A; p.(Gly899Arg)	chr21:g.44087854G>A	-	-	-	-	0.044	-0.97	0.202	0.058	0.08
Family 4 (HOM)	c.3222C>A; p.(Cys1074*)	chr21:g.44098410G>T	-	-	-	-	NA				

Table 3.3: Biallelic *TRAPPC10* variants identified in affected individuals from four families

Grey shading indicates that the pathogenicity of the *TRAPPC10* variants were not confirmed in this family. Abbreviations: HOM; homozygous, AF; allele frequency, NA; not applicable, pat; paternally inherited, mat; maternally inherited.

A third family from Spain was identified through GeneMatcher [244], comprising a 6 year old affected female with non-syndromic severe intellectual disability and autism (Table 3.2). In this family, trio WES identified compound heterozygous missense variants in *TRAPPC10*, involving a paternally inherited chr21:g.44079598C>T; NM_003274.4: c.1504C>T; p.(Leu502Phe) [hg38] and two maternal variants inherited *in cis* (chr21:g.44083081G>A; NM_003274.4: c.2017G>A; p.(Ala673Thr) and chr21:g.44087854G>A; NM_003274.4: c.2695G>A; p.(Gly899Arg) [hg38]) (Figure 3.4). All three variants are listed at low frequency in gnomAD v2.1.1/v3.1.2 and have mixed predictions for missense pathogenicity (Table 3.3). The maternal (p.(Ala673Thr)/p.(Gly899Arg)) variant *in cis* haplotype was hypothesised to be acting as a pathogenic allele. We undertook studies to assess the function and interaction of *TRAPPC10* variants identified in these three families.
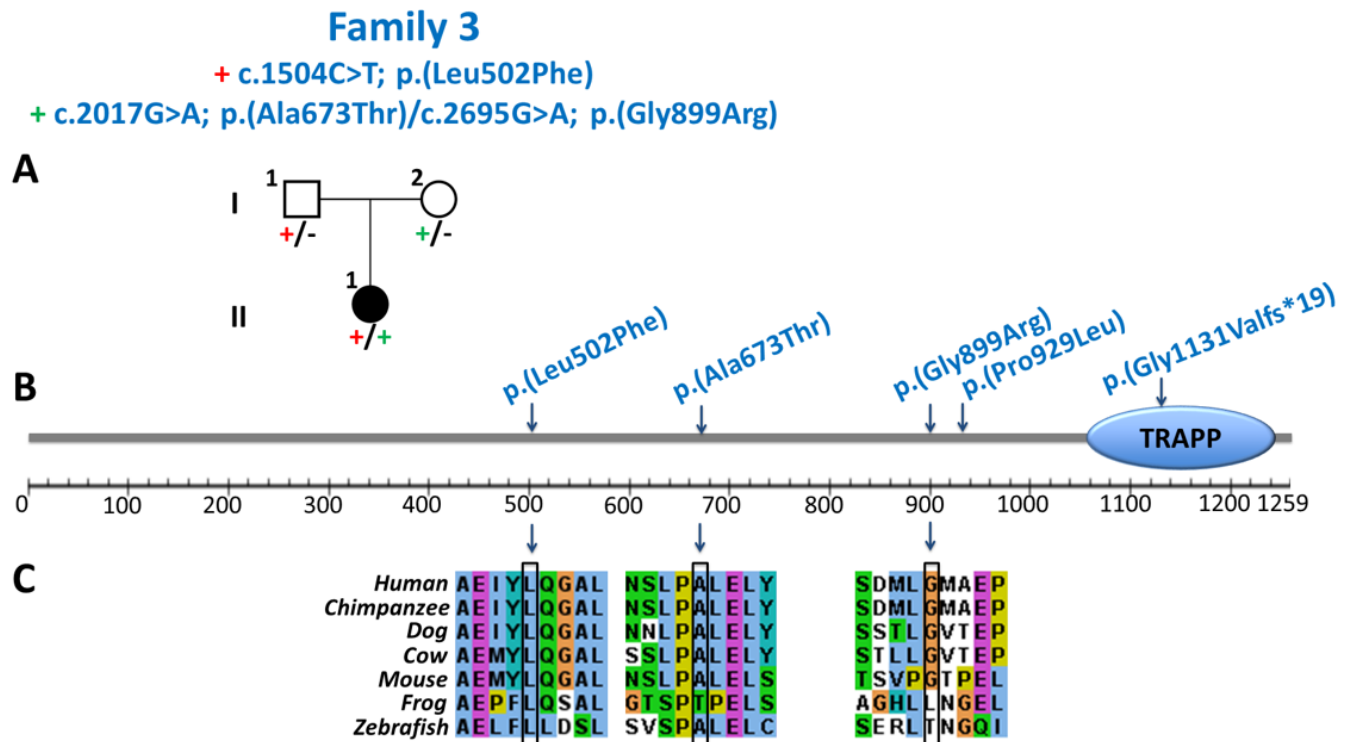
Figure 3.4: *TRAPPC10* variants identified in Family 3

(**A**) Family 3: Nuclear family with a single affected individual compound heterozygous for *TRAPPC10* variants; paternally inherited c.1504C>T; p.(Leu502Phe) variant ('+') and maternally inherited haplotype with c.2017G>A; p.(Ala673Thr) and c.2695G>A; p.(Gly899Arg) variants ('+') with confirmed co-segregation ('-' indicates the wild type allele). (**B**) Schematic diagram of the TRAPPC10 protein showing three *TRAPPC10* missense variants identified in Family 3 alongside pathogenic variants identified in Families 1 and 2. (**C**) Multi-species alignment showing conservation of the molecular region encompassing the p.(Leu502Phe), p.(Ala673Thr) and p.(Gly899Arg) variants.

3.4.2 Cellular/molecular studies

3.4.2.1 Confirmation of FLAG-tagged *TRAPPC10* variants

FLAG-tagged pDONR™201 constructs of p.Gly1131Valfs*19, p.Pro929Leu, p.Leu502Phe, p.Ala673Thr, p.Gly899Arg and combined p.Ala673Thr/p.Gly899Arg variants were created for use in subsequent molecular studies to learn more about the subcellular outcomes resulting from these *TRAPPC10* variants. FLAG-tagged pDONR™201 constructs were confirmed with BsrgI restriction enzyme digestion with the expected resultant product size and then transfected into HeLa cells and confirmed by Western blot analysis (Figure 3.5). The commercially available anti-TRAPPC10 antibody (Santa Cruz) is directed to the C-terminal 1162-1258 amino acids downstream of the p.Gly1131Valfs*19 variant, so as expected we did not detect TRAPPC10 for this variant due to truncation of the protein. Presence of the p.Gly1131Valfs*19 mutant protein was confirmed using anti-FLAG antibody confirming a smaller size band (~130kDa) due to the truncated TRAPPC10 molecule (Figure 3.5).
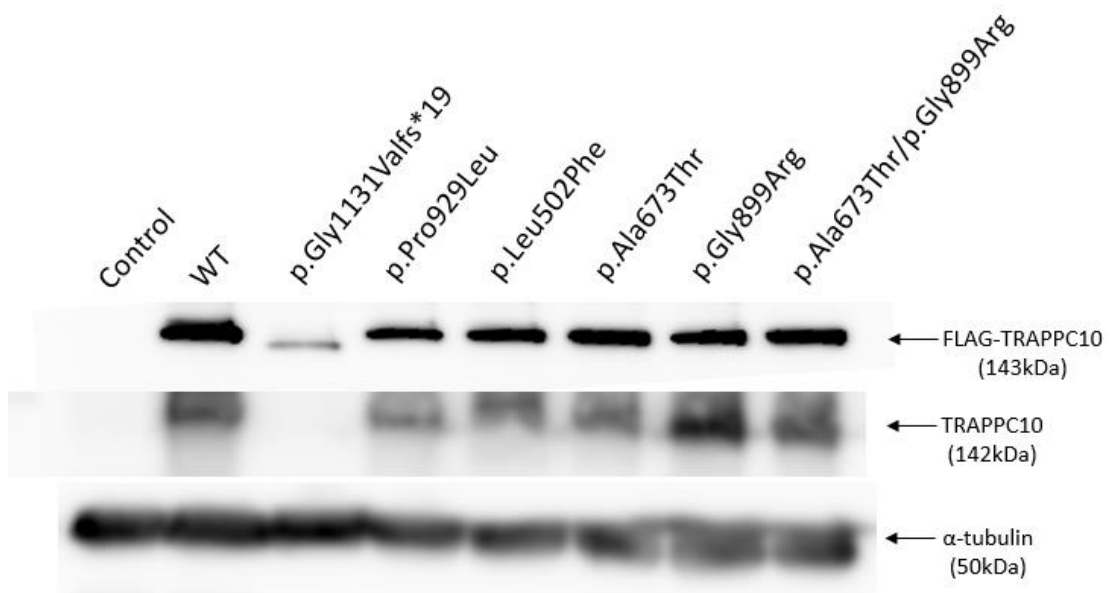
Figure 3.5: Characterisation of FLAG-tagged constructs in HeLa cells

Lysate was prepared from HeLa cells transfected with FLAG-tagged wild type (WT) TRAPPC10, p.Gly1131Valfs*19, p.Pro929Leu, p.Leu502Phe, p.Ala673Thr, p.Gly899Arg and combined p.Ala673Thr/p.Gly899Arg variants, then fractionated by SDS-PAGE and probed for FLAG and TRAPPC10. Tubulin was included as a loading control. Parental HeLa cells were included as a control.

3.4.2.2 Confirmation of HAP1 *TRAPPC10^{-/-}* cell lines and rescue studies with wild type TRAPPC10 transfection

HAP1 *TRAPPC10^{-/-}* cell lines were created for use in subsequent studies to assess the cellular effects of a complete absence of TRAPPC10 protein and to assess the molecular outcome of *TRAPPC10* variants. HAP1 cells were used for initial knockout cell studies as they are haploid and easier to transfect and study than many other cell lines, including HEK293 cells, which were used in later knockout cell studies. Two HAP1 *TRAPPC10^{-/-}* cell lines (G2-19 and G2-21) were selected due to an abnormal clustering appearance compared with HAP1 parental cells (Figure 3.6A-F). Knockout verification was performed with Western blotting showing an absence of detectable TRAPPC10 (Figure 3.6J). *TRAPPC10* knockout was further confirmed in HAP1 cell lines by dideoxy sequencing of PCR products using primers designed around the target region for guide 2 RNA in exon 3 of *TRAPPC10* (Appendix 1, Table 1.1). A 27bp insertion (TCACAAAAAAAACGTTTTTTTTTTTTT) was identified in the HAP1 G2-21 cell line. While likely in frame, the insertion of 27bp (encoding nine amino acids) is considered likely to result in loss of function of TRAPPC10. Unfortunately, no PCR product could be generated in G2-19 cells despite primer redesign and optimisation (Figure 3.7).
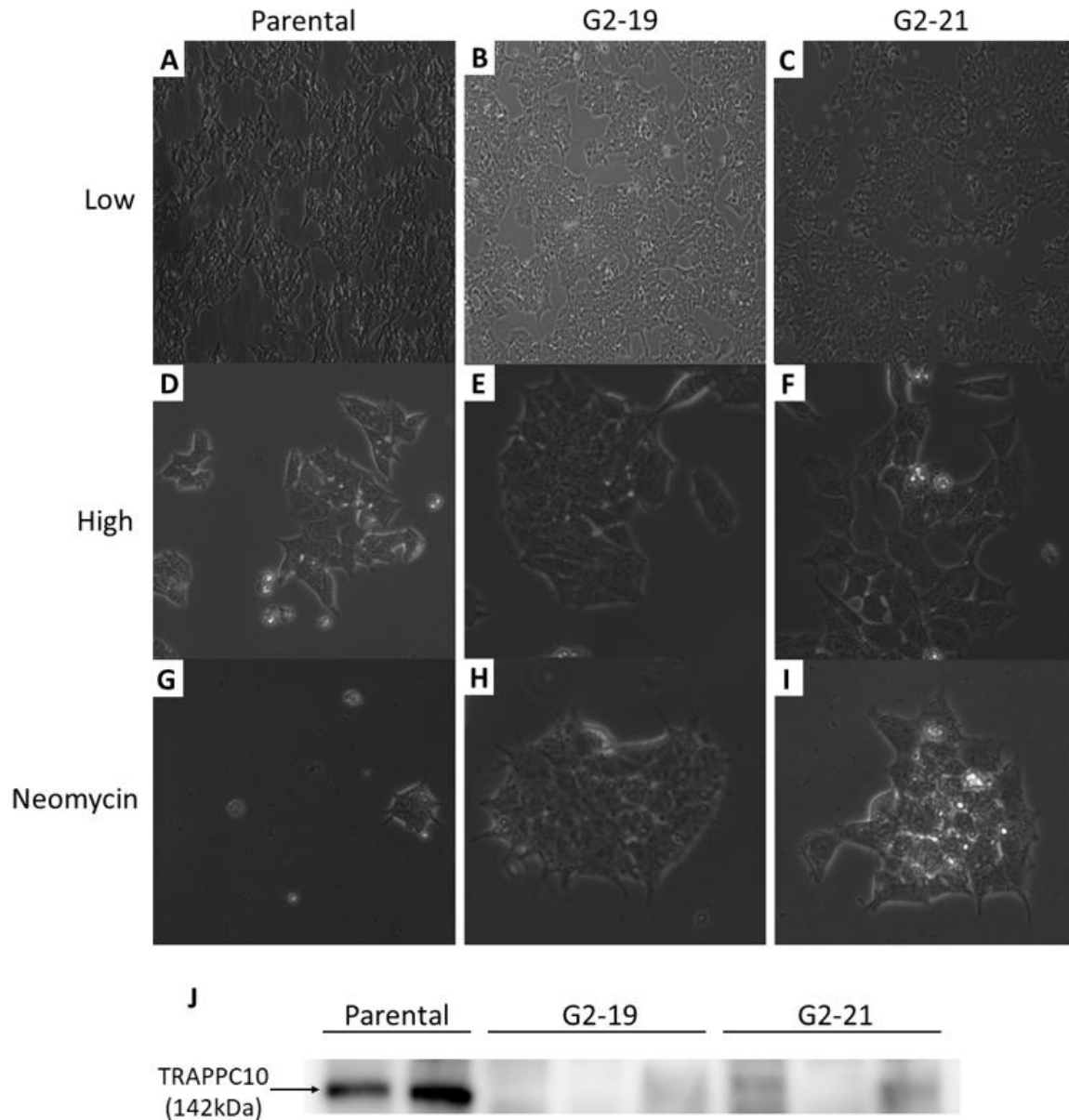
Figure 3.6: Microscopic appearance and confirmation of HAP1 *TRAPPC10*⁻/⁻ cell lines

(**A,D**) Light microscopy at low (100x) and high (400x) power of HAP1 parental cells demonstrating the normal appearance. (**B,E**) Appearance of G2-19 HAP1 *TRAPPC10*⁻/⁻ cells showing abnormal clustering. (**C,F**) Appearance of G2-21 HAP1 *TRAPPC10*⁻/⁻ cells showing abnormal clustering. (**G-I**) The appearance of the cells after neomycin selection following transfection of the G2-19 and G2-21 cell lines with FLAG-tagged WT TRAPPC10 constructs showing no change in clustering. (**J**) Western blot with probe for TRAPPC10 of cell lysates from HAP1 parental, G2-19 and G2-21 HAP1 *TRAPPC10*⁻/⁻ cells with exposure for 3 minutes, confirming the absence of TRAPPC10 in both *TRAPPC10*⁻/⁻ cell lines.
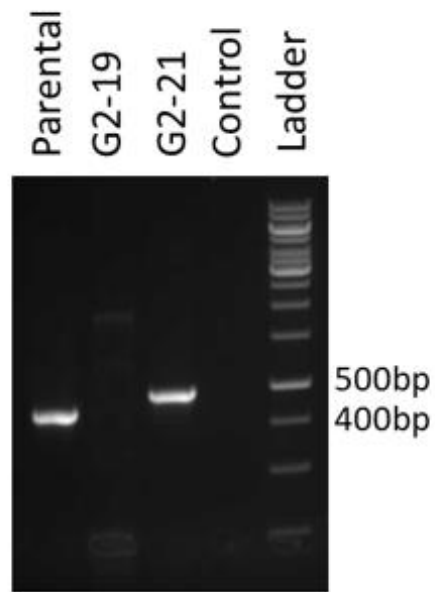
Figure 3.7: Gel electrophoresis of PCR product from HAP1 *TRAPPC10*[-/-] cells

A larger band was detected in G2-21 *TRAPPC10*[-/-] cells compared with HAP1 parental cells consistent with the insertion of the 27bp fragment confirmed by dideoxy sequencing. No product was detectable for the G2-19 cell line. A 1Kb plus ladder was used.

To confirm that the abnormal clustering phenotype observed with the G2-19 and G2-21 HAP1 *TRAPPC10$^{-/-}$* cell lines was due to an absence of TRAPPC10, rescue studies attempted to transfect WT TRAPPC10 into the *TRAPPC10$^{-/-}$* cell lines, and observe whether the cell clustering phenotype remained. G2-19 and G2-21 HAP1 *TRAPPC10$^{-/-}$* cell lines were cotransfected with plasmids containing both C-terminus red fluorescent protein (mRFP-C1) and FLAG-tagged WT TRAPPC10 pDONR™201 constructs to create RFP-labelled cells that can be easily visualised under a fluorescent microscope, with transfection of WT TRAPPC10 confirmed through the presence of FLAG. These cells could then be selected and observed for changes in clustering properties, although unfortunately this was not possible as the cells did not survive cotransfection. Both HAP1 *TRAPPC10$^{-/-}$* cell lines were then successfully transfected with only FLAG-tagged WT TRAPPC10, followed by neomycin selection to determine whether this rescued the abnormal cell clustering phenotype. There were no observed changes in the microscopic appearance of the cells following WT TRAPPC10 transfection (Figure 3.6G-I) with no rescue of the abnormal phenotype.

3.4.2.3 *TRAPPC10* variants exhibit reduced interaction with TRAPPC2L, destabilise the TRAPP II complex and result in an anterograde trafficking defect

We sought to determine the molecular consequences of each *TRAPPC10* genetic variant. As we previously determined that TRAPPC10 interacts with TRAPPC2L in the TRAPP II complex [276], we used yeast two-hybrid assays to investigate how TRAPPC10 alteration affects TRAPPC2L interactions. When TRAPPC10 was in the bait (pGBKT7) or prey (pGADT7) vector, TRAPPC2L interaction was observed (Figure 3.8A,B). However, both p.Gly1131Valfs*19 and p.Pro929Leu variants showed reduced TRAPPC2L interaction (Figure 3.8A,B), though the proteins were expressed in the yeast cells (not shown). The p.Leu502Phe, p.Ala673Thr, p.Gly899Arg and combined p.Ala673Thr/p.Gly899Arg variants showed no difference in interaction with TRAPPC2L compared to WT TRAPPC10 (not shown).

This was confirmed quantitatively by measuring β-galactosidase activity (Figure 3.8C), one of the yeast two-hybrid system markers. Given this, we generated a patient lymphoblastoid cell line (LCL) (1-IV:2; *TRAPPC10* p.(Gly1131Valfs*19)) to investigate TRAPP complex outcomes. We examined core TRAPP protein (TRAPPC2, TRAPPC2L, TRAPPC3) levels, as well as TRAPP III-specific (TRAPPC8,

TRAPPC12), and TRAPP II-specific (TRAPPC10, TRAPPC9) levels. TRAPPC2L, TRAPPC2 and TRAPPC3 levels were slightly reduced in the p.(Gly1131Valfs*19) lysate compared to control (Figure 3.8D). As expected we did not detect TRAPPC10 as the commercially available TRAPPC10 antibody is directed to the molecular region downstream (C-terminal) to the p.(Gly1131Valfs*19) alteration. Intriguingly, we also noted an absence of full length TRAPPC9, although in contrast, TRAPP III-specific proteins levels were not affected. We then fractionated p.(Gly1131Valfs*19) and control LCL lysates on a size exclusion column. While TRAPPC10 and TRAPPC9 fractionated as expected in controls, they were absent from p.(Gly1131Valfs*19) lysate (Figure 3.8D) indicating an absence of TRAPP II (peak in fraction 28 of control cells). There was no change in core protein (TRAPPC2, TRAPPC2L, TRAPPC3) fractionation consistent with their presence in TRAPP III, although the portion that co-fractionates with TRAPP II was reduced. Similarly, no change in TRAPP III proteins (TRAPPC8, TRAPPC12) fractionation was seen, indicating that this complex was not affected in p.(Gly1131Valfs*19) cells.
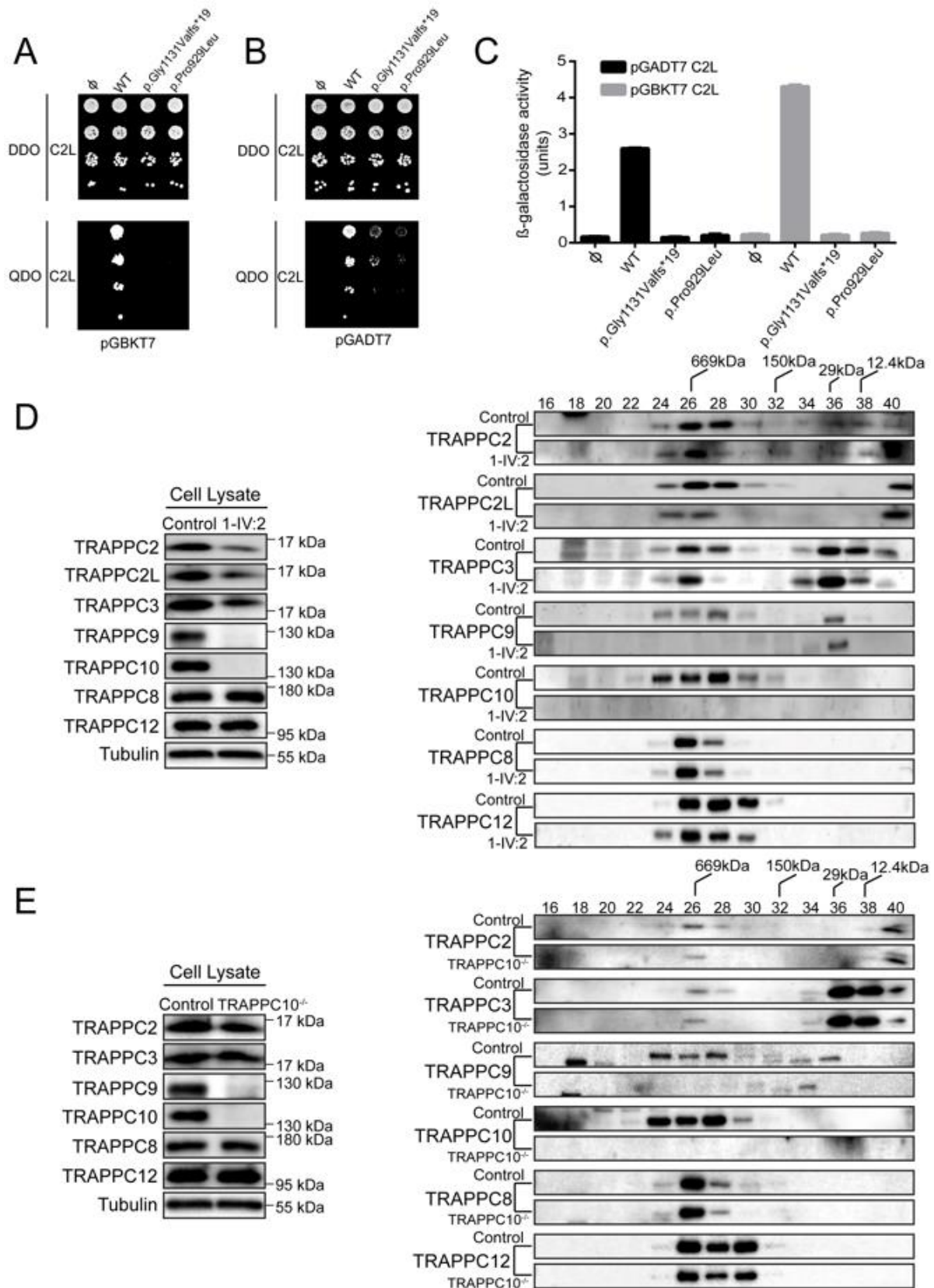
Figure 3.8: TRAPP II complex integrity is affected by *TRAPPC10* variants and in *TRAPPC10^-/-* cells

Yeast cells transformed with pGBKT7 or pGBKT7 expressing either wild type TRAPPC10, the p.Gly1131Valfs*19 variant or the p.Pro929Leu variant (**A**), or pGADT7 or pGADT7 expressing either wild type TRAPPC10, the p.Gly1131Valfs*19 variant or the p.Pro929Leu variant (**B**) were mated with yeast expressing either pGBKT7-TRAPPC2L (**A**) or pGADT7-TRAPPC2L (**B**). Serial dilutions of the diploid cells were grown on SC medium lacking leucine and tryptophan (DDO) or SC medium lacking leucine, tryptophan, histidine and adenine (QDO). (**C**) Diploid cells from (**A**) and (**B**) were quantitatively tested for β-galactosidase activity. Units of activity were calculated according to the following formula: (OD420 x 1000)/(OD600 x time in hours). Note that western analysis revealed that all forms of TRAPPC10 and the two variants were expressed in the yeast cells. (**D**) Lymphoblastoid cells from control or an individual homozygous for the *TRAPPC10* p.(Gly1131Valfs*19) variant were lysed and probed for the TRAPP proteins indicated and for tubulin as a representative loading control. The lysates were fractionated on a Superose 6 size exclusion column. Fractions of 0.5ml were collected. Neighbouring fractions were pooled and fractionated by SDS-PAGE and probed for the indicated TRAPP proteins. (**E**) Wild type HEK293 or *TRAPPC10$^{-/-}$* cells were lysed and probed for the TRAPP proteins indicated and for tubulin as a representative loading control. The lysates were fractionated on a Superose 6 size exclusion column. Fractions of 0.5ml were collected. Neighbouring fractions were pooled and fractionated by SDS-PAGE and probed for the indicated TRAPP proteins. The fractionation of molecular size standards are indicated above the top-most panel of the size exclusion portion of (**D**) and (**E**). Data courtesy of Professor Michael Sacher.

We then examined the consequences of TRAPPC10 loss on membrane trafficking. Since we were unable to perform trafficking assays in smaller LCLs, we generated HEK293 TRAPPC10 knockout cell lines using CRISPR/Cas9. Three different *TRAPPC10* knockout lines were generated (Figure 3.9), two were examined and behaved identically (*TRAPPC10$^{-/-}$* KO1/KO7); as such we present data from one line (KO7, hereafter referred to as *TRAPPC10$^{-/-}$*).

We first examined TRAPP protein levels in *TRAPPC10$^{-/-}$* cells. Consistent with LCLs observations, we noted a slight reduction in core TRAPP proteins, no effect on TRAPP III-specific proteins (TRAPPC8, TRAPPC12), and the absence of TRAPPC9, consistent with size exclusion chromatography results (Figure 3.8E). We then performed the VSVG-GFP ts045 trafficking assay [323] in *TRAPPC10$^{-/-}$* cells. In this assay, the marker protein VSVG-GFP ts045 is retained in the ER at elevated temperature, but synchronously released at 32°C (Appendix 2, Figure A2.2). The fluorescence-tagged marker protein accumulated in the Golgi region in controls, with peak signal at ~25 minutes (Figure 3.10A, black curve).
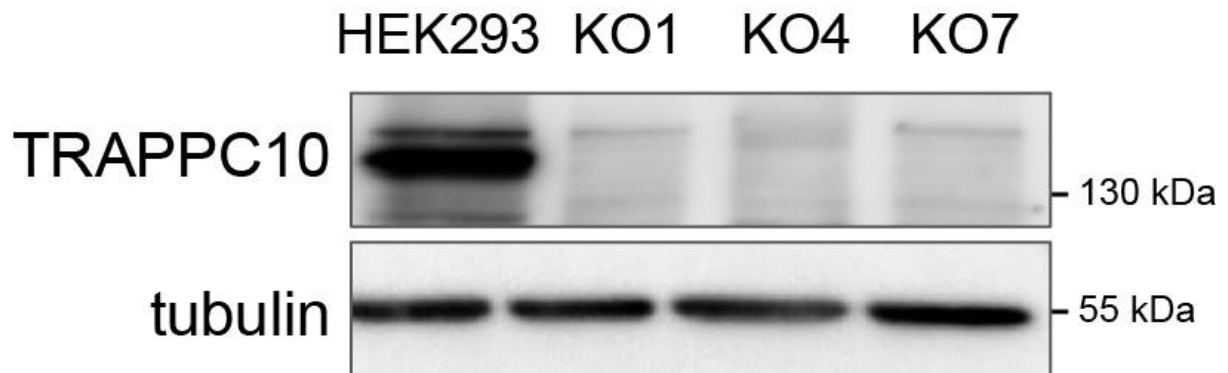
Figure 3.9: Characterisation of the *TRAPPC10⁻ᐟ⁻* cell lines in HEK293 cells

Lysate was prepared from parental HEK293 cells and from 3 clones (KO1, KO3 and KO7) that were treated with sgRNA targeting *TRAPPC10*. The lysates were fractionated by SDS-PAGE and probed for TRAPPC10. Tubulin was included as a loading control. Data courtesy of Professor Michael Sacher.
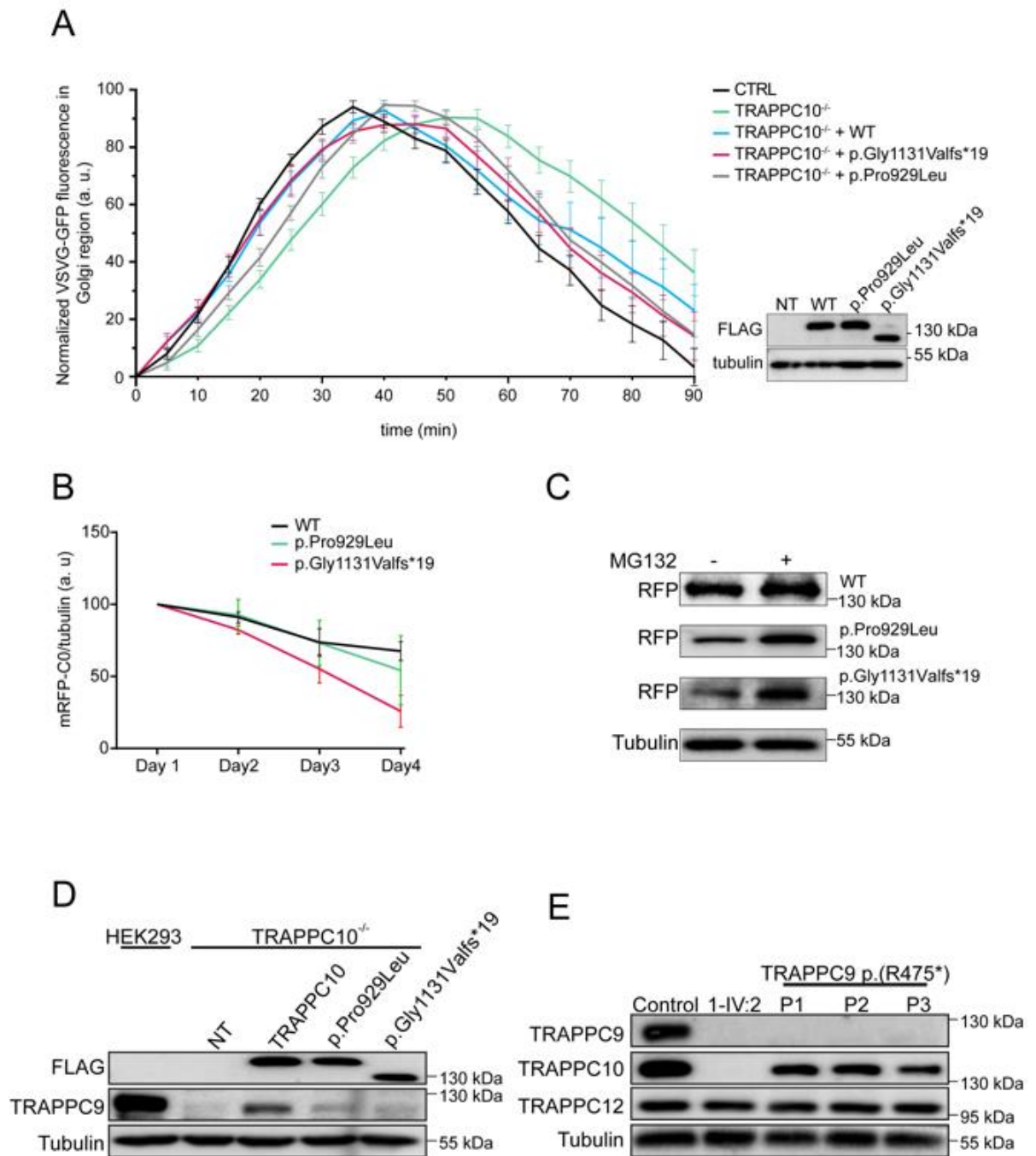
Figure 3.10: Cells devoid of TRAPPC10 display a membrane trafficking defect and lack detectable levels of TRAPPC9

(**A**) Wild type HEK293 cells or *TRAPPC10^-/-^* cells either not transfected or transfected with FLAG-tagged wild type TRAPPC10 or one of the *TRAPPC10* variants indicated were infected with VSVG-GFP ts045 4 hours after transfection. After an overnight incubation at 40°C, the cells were shifted to 32°C and imaged every minute. The fluorescence intensity in the Golgi was quantified and plotted versus time as described in the materials and methods section. The inset shows a western blot for the transfected proteins probed with anti-FLAG antibody. Representative images are shown in Appendix 2, Figure A2.2.

136

(**B**) RFP-tagged versions of either wild type TRAPPC10, the p.Gly1131Valfs*19 variant or the p.Pro929Leu variant were transfected into HeLa cells and lysates prepared every day for 4 days. The lysates were fractionated by SDS-PAGE and probed with anti-mCherry antibody to reveal the RFP-tagged construct. Though the trends were consistent, statistical significance was seen only at day 4. The half-life for each overexpressed protein was found to be approximately 5.4 days, 4.4 days and 3.1 days for the wild type, p.Gly1131Valfs*19 and p.Pro929Leu variants, respectively. (**C**) RFP-tagged TRAPPC10 and the p.Gly1131Valfs*19 and p.Pro929Leu variants were transfected into HEK293 cells. The cells were either untreated (-) or treated (+) with 50 μm MG132 for 24 hours on the third day post-transfection. Lysates were then prepared, fractionated by SDS-PAGE and probed with anti-mCherry and tubulin as a loading control. (**D**) *TRAPPC10$^{-/-}$* cells were either untransfected or transfected with FLAG-tagged wild type TRAPPC10, the p.Gly1131Valfs*19 variant or the p.Pro929Leu variant. After 24 hours, lysates were prepared and probed for the FLAG constructs to verify expression, TRAPPC9 and tubulin as a loading control. Parental HEK293 cells were also probed to assess the level of TRAPPC9 in the presence of TRAPPC10. (**E**) Lysates were prepared from lymphoblastoid cells from either control, an individual homozygous for the TRAPPC10 p.(Gly1131Valfs*19) alteration or three individuals homozygous for the *TRAPPC9* p.(Arg475*) variant and subjected to western analysis to reveal the proteins indicated. Data courtesy of Professor Michael Sacher.

From that point onward the signal was cleared from the Golgi due to the protein being transported to the plasma membrane. In contrast, the arrival into, trafficking through and release from the Golgi of the marker protein in *TRAPPC10*^{-/-} cells was delayed compared to controls (Figure 3.10A, green curve). Importantly, the trafficking delay was shown to be TRAPPC10-dependent as cell transfection with WT TRAPPC10 resulted in a trafficking defect rescue (Figure 3.10A, blue curve). Although both p.Gly1131Valfs*19 and p.Pro929Leu variants were able to partially rescue these defects, neither rescued anterograde trafficking to the same extent as WT protein (Figure 3.10A, red/grey curves). Taken together, these results demonstrate that the *TRAPPC10* variants cause partial functional defects, and that the absence of TRAPPC10 results in a concomitant absence of TRAPPC9.

3.4.2.4 Expression of TRAPPC10 is important for the detection of a cellular pool of TRAPPC9

As no antibody was available to determine whether mutant TRAPPC10 p.(Gly1131Valfs*19) polypeptide may be expressed in affected individuals, we generated WT and both (p.Gly1131Valfs*19, p.Pro929Leu) constructs to investigate protein stability. The constructs were transfected into HEK293 cells and lysates prepared daily for four days. As shown in Figure 3.10B, WT protein levels on day 4 dropped to ~70% of that on day 1. In contrast, levels of both altered TRAPPC10 proteins (day 4) dropped to 25-50% (day 1), with the drop in the p.Gly1131Valfs*19 protein being most dramatic. The decrease in expression of WT and both p.Gly1131Valfs*19 and p.Pro929Leu TRAPPC10 proteins was blocked by inclusion of proteasome inhibitor MG132 (Figure 3.10C). These results suggest that the two mutant proteins are subject to enhanced degradation compared to WT protein.

The absence of TRAPPC9 in TRAPPC10 p.(Gly1131Valfs*19) LCLs prompted us to closely examine their inter-relationship. We first investigated whether TRAPPC10 expression in *TRAPPC10*^{-/-} cells would restore TRAPPC9 levels. Upon transfection of WT TRAPPC10 into *TRAPPC10*^{-/-} cells, the appearance of TRAPPC9 was noted, although not to parental HEK293 cell levels (Figure 3.10D). Overexpression of the p.Pro929Leu variant, and to a much lesser degree p.Gly1131Valfs*19, restored some TRAPPC9 protein although again not to the same extent as WT TRAPPC10.

We next investigated whether a reciprocal absence of TRAPPC10 occurred in cells devoid of TRAPPC9. We probed LCL lysates from three individuals with the p.(Arg475*) TRAPPC9 variant [293]. As expected TRAPPC9 was not detected in these cells (Figure 3.10E,F), and nor was it detected in TRAPPC10 p.(Gly1131Valfs*19) cells (as shown above). Importantly, while TRAPPC10 levels were significantly reduced in p.(Arg475*) TRAPPC9 lysates, the protein was nevertheless present. Collectively, these results suggest that the detection of a cellular pool of TRAPPC9 is strongly dependent upon the expression of full length TRAPPC10.

3.4.3 Mouse studies

3.4.3.1 *Trappc10*$^{-/-}$ mice have similar neurodevelopmental deficits to patients with biallelic *TRAPPC10* variants

We evaluated tissue sections and datasets from a *Trappc10*$^{tm1b(EUCOMM)Wtsi}$ mouse model, with a focus on the phenotypical components of the human disorder in particular neuroanatomical findings. At weaning age, mouse survival was evaluated from successfully genotyped mice originating from multiple litters derived from a heterozygous-by-heterozygous breeding scheme. We obtained the expected number of WT, heterozygous and homozygous mice. Male and female mice were weighed the same day each week from 4 until 16 weeks of age (Appendix 2, Figure A2.3).

Using a recently developed robust approach to assess 63 brain parameters across 23 brain regions [324], we analysed neuroanatomical defects in adult *Trappc10*$^{-/-}$ mice blinded for genotype. To minimize environmental and genetic variation, male mice aged 16 weeks were used. In the homozygous mutant mice, many neuroanatomical parameters were reduced in size when compared to WTs (Figure 3.11A,B). The total brain area parameter was significantly reduced (14%, P=0.0003), concomitantly with smaller white matter structure size including the genu of the corpus callosum (-25%, P=0.026), hippocampal commissure (-38%, P=0.013), hippocampal fimbria (-17%, P=0.0049), anterior commissure (-28%, P=3.8E10-8) and the internal capsule (-23%, P=1.1E10-6) (Figure 3.11C,D).
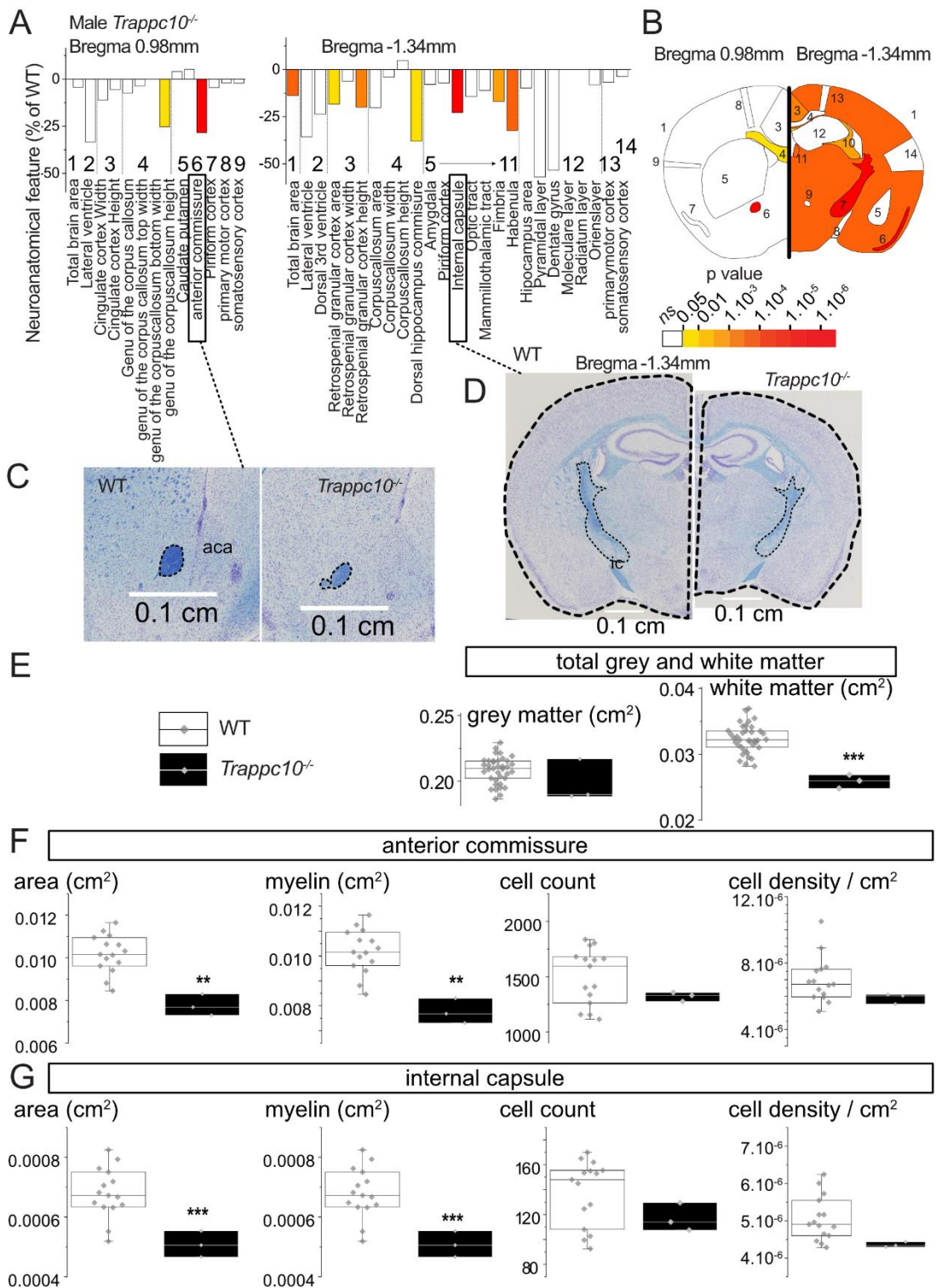
Figure 3.11: Mouse studies implicate TRAPPC10 in the formation of white matter structures

(**A**) Histograms for three homozygous *Trappc10⁻ᐟ⁻* mice showing variation (decreased-minus scale or increased-positive scale) in areas and lengths expressed as percentage of 498 WTs together with a colour map indicating the significance level (white indicates not significant). The list of measurements and corresponding numbers are shown below the histograms. (**B**) Schematic representation of a section at Bregma +0.98 mm and Bregma -1.34 mm. Coloured regions indicate the presence of at least one significant parameter within the brain region at the 0.05 level. (**C-D**) Nissl-stained sagittal brain sections from Trappc10[tm1b/tm1b] mice, showing the anterior commissure (**C**) and the total brain area and internal capsule size reduction in mutants (**D**). The corresponding scale is shown in each panel. (**E**) Box plots of combined grey and white matter structures expressed in cm^2 using a set of 37 local WTs (same genetic background, housing conditions, age, sex, experimental and necropsy dates) compared to three Trappc10[tm1b/tm1b] mice. (**F-G**) Box plots with raw data points detailing the cellular and myelination characterisation of the anterior commissure (**F**) and internal capsule (**G**) using a set of 15 local WTs. Statistical analyses were performed with GraphPad Prism 8.0.2, using two-tailed Student's t-tests of equal variances. *$p<0.05$ **$p<0.01$ ***$p<0.001$. Data courtesy of Dr Binnaz Yelcin.

When combining neuroanatomical parameters from either grey or white matter structures, only white matter structures were reduced in size (-20%, P=8E10-6) (Figure 3.11E). Interestingly, the size reductions of the anterior commissure and internal capsule were correlated with loss of myelination whilst oligodendrocyte cell population count and density were unaffected (Figure 3.11F,G). Together these findings suggest that both mouse and human microcephaly stems from defects in white matter structures and myelin biogenesis.

Given the short stature exhibited by patients with *TRAPPC10*-related disorder we also evaluated *Trappc10$^{-/-}$* mice growth parameters. While long bone length was marginally reduced in both sexes, it only reached significance for female *Humerus* (Appendix 2, Figure A2.3A), although body length overall was unaffected (Figure A2.3B). While weight was unaffected in *Trappc10$^{-/-}$* males, female *Trappc10$^{-/-}$* mice were overweight, showing a rapid weight increase commencing week 10 (Figure A2.3C-E). We investigated the cause of the female obesity phenotype and noted a highly significant fat mass increase (P=9E-08; Figure A2.3D), alongside a blood cholesterol increase (P=4E-05) and high-density lipoprotein levels (P=0.003).

We finally sought to further characterise the phenotypical relationship between *Trappc10$^{-/-}$* mice, and a previously published *Trappc9$^{-/-}$* mouse model [303]. Due to the unavailability of live *Trappc9$^{-/-}$* animals, these evaluations were limited to assessments undertaken on historical brain tissue sections. Overall, the impact of TRAPPC10 loss on brain anatomy was more severe than TRAPPC9 loss (Appendix 2, Figure A2.4). While there were some overlaps between affected brain regions (e.g. corpus callosum reduced size), some neuroanatomical phenotypes were specific to *Trappc10$^{-/-}$* including reduced size of the anterior commissure.

3.4.4 Identification of a further family with biallelic *TRAPPC10* variants

Since our original studies were completed and published [325], we recently identified a further individual affected by *TRAPPC10*-related disorder through GeneMatcher [244], for whom we have unfortunately not had an opportunity to perform functional studies. This family (Family 4) comprises a consanguineous Iranian family with a single female child affected with a microcephalic (-6.77SD) neurodevelopmental disorder (Table 3.2) characterised by profound intellectual disability and developmental delay (non-verbal, never achieved walking although is able to sit

independently), short stature, occasional tonic-clonic seizures, involuntary dystonic movements, autistic features, dysmorphic facial features and bilateral sensorineural hearing loss. Trio exome sequencing identified a single candidate segregating homozygous variant as the likely cause of neurodevelopmental disorder in exon 21/23 of *TRAPPC10* (chr21:g.44098410G>T; NM_003274.4: c.3222C>A; p.(Cys1074*)) [hg38]) (Figure 3.12A,B) located within the C-terminal TRAPP domain (Figure 3.12C), that is predicted to undergo NMD and is absent from gnomAD v2.1.1/v3.1.2 (Table 3.3). A pathogenic homozygous *GJB2* variant was also identified as the likely cause of deafness in this individual.
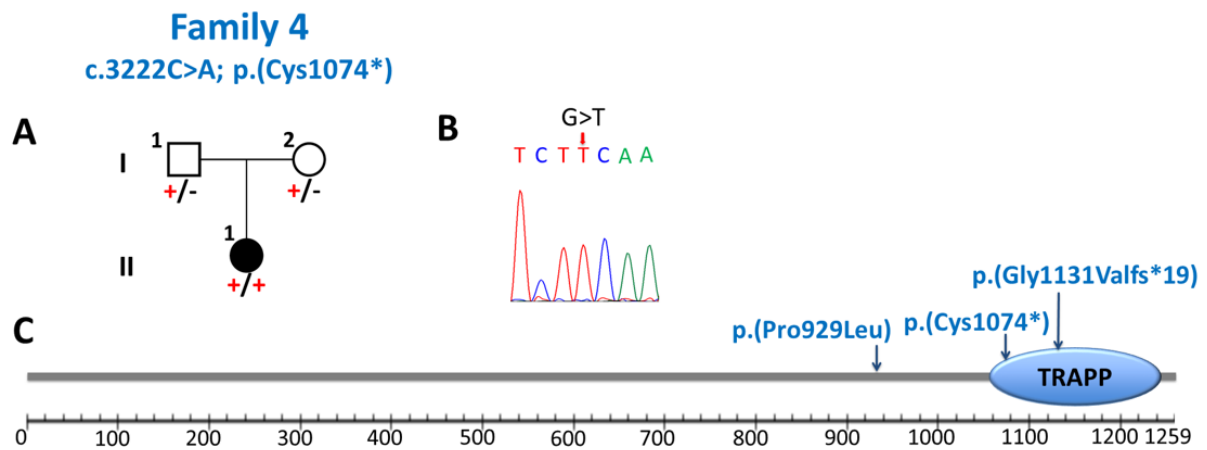
Figure 3.12: Pedigree and genetic studies for an individual subsequently identified with biallelic *TRAPPC10* variants

(**A**) Family 4: Nuclear family with a single affected individual homozygous for the *TRAPPC10* c.3222C>A; p.(Cys1074*) variant ('+') with confirmed co-segregation ('-' indicates the wild type allele). (**B**) Dideoxy sequence chromatogram of the affected individual homozygous for the *TRAPPC10* c.3222C>A; p.(Cys1074*) variant. (**C**) Schematic diagram of the TRAPPC10 protein identifying all three pathogenic *TRAPPC10* sequence variants and the TRAPP domain.

3.4.5 Discovery of three novel and consolidation of two recently described TRAPPopathy disorders

Following on from these studies of *TRAPPC10*-related disorder, further investigation was undertaken to identify whether variants in other TRAPP complex subunits may be associated with human disease. This involved extensive literature review and data analysis from DDD CAP#303 and the 100,000 genomes project, described in Appendix 3.

Three individuals with biallelic *TRAPPC1* variants (p.(Val121Alafs*3), p.(22_24del), p.(His98Pro), p.(His72Arg) and c.170+5G>A) had a severe microcephalic neurodevelopmental disorder with seizures and significant motor impairment. A single individual with putative pathogenic compound heterozygous *TRAPPC5* variants p.(Cys139Serfs*33) (low AF 0.00002627 in gnomAD with no homozygotes) and p.(Gly99Val) (within the TRAPP domain, AF 0.0004007 in gnomAD with no homozygotes, amino acid conservation to zebrafish and multiple *in silico* tools predicting pathogenicity) had seizures and other features consistent with other TRAPPopathies, including microcephaly, severe neurocognitive impairment, seizures, choreoathetosis, cerebral visual impairment/blindness and neurogenic bladder. Additionally, three singleton individuals displayed overlapping clinical features including microcephaly, short stature and intellectual disability, all associated with biallelic *TRAPPC8* variants (p.(Ser979del), p.(Asn1340Ser), p.(Leu542Val) and p.(Leu31Phe)) (see Appendix 3, A3.3.2).

Two further patients were identified with likely *TRAPPC3*-related disorder (associated with p.(Ile51Thr) and p.(Leu131Phe) variants) and a strikingly similar BBS-like phenotype to that seen in a single previously published family, consolidating this as a likely ciliopathy disease gene [281]. Additionally, two patients were identified with biallelic *TRAPPC14* variants (p.(Arg364Cys) and p.(Glu465Lys)), associated with a possibly concordant phenotype to that in the only other candidate family previously published with this condition [320] (see Appendix 3, A3.3.3).

Together the results of these studies provide evidence supporting the discovery of three further TRAPP complex subunits as candidate TRAPPopathy disorder genes (*TRAPPC1*, *TRAPPC5* and *TRAPPC8*). These studies also provide new evidence to consolidate two TRAPP complex subunits (*TRAPPC3* and *TRAPPC14*), each

previously described in only a single family, as likely causative of neurodevelopmental TRAPPopathies.

3.5 Discussion

This study defines biallelic *TRAPPC10* variants as a cause of a neurodevelopmental TRAPPopathy disorder, with core clinical features including microcephaly, severe global developmental delay and intellectual disability, short stature and pervasive behavioural abnormalities (Table 3.2). Additional variable features include hypotonia, gait abnormalities and seizures. The overlapping phenotypical nature of the TRAPPopathies suggests a common pathomolecular basis of disease involving disruption of TRAPP II/III complex functions. The clinical features of *TRAPPC10*-related disorder also show extensive overlap with many other Golgipathy disorders [7, 106] indicative of a common disease mechanism relating to disrupted Golgi trafficking processes. We observed notable phenotypic overlap with *Trappc10$^{-/-}$* mice which display microcephaly, reduced size of white matter brain structures with hypomyelination and skeletal involvement. Brain abnormalities in mice appear to be more severe than those in humans, in which neuroimaging identified profound thinning of the corpus callosum but no other white matter abnormalities or hypomyelination. However, as neuroimaging was only available on a single affected individual who appears to display the least severe phenotype of those assessed, it is possible that more extensive abnormalities may be present in other affected individuals who display a greater degree of microcephaly.

Our genetic and clinical studies in two families (Families 1 and 2) in combination with our functional investigations on both *TRAPPC10$^{-/-}$* and patient-derived cells indicate that TRAPPC10 loss of function likely underlies the clinical phenotype. The poorly concordant phenotype, *in silico* predictions and functional studies of the *TRAPPC10* missense variants identified in Family 3 suggest that these variants are likely benign and that this individual does not have *TRAPPC10*-related disorder. However while molecular studies remain to be undertaken, the subsequent identification of a further affected individual (Family 4) with a closely overlapping clinical phenotype and homozygous pathogenic loss of function *TRAPPC10* p.(C1074*) variant identifies a third family with *TRAPPC10*-related disorder (Figure 3.12).

Previous subcellular trafficking studies of TRAPPC10 depleted COS-7 [266] and HEK293 [267] cells determined that TRAPPC10 mediates early Golgi (although not ER) trafficking, revealing disrupted Golgi architecture and vesicular accumulation indicative of an anterograde trafficking defect. Our cell model membrane trafficking assays are in complete alignment with such a TRAPPC10 molecular role, defining delayed Golgi trafficking associated with gene KO which importantly could be rescued, though not completely, by WT, and to a lesser extent by mutant TRAPPC10. These findings may in part be explained by differing outcomes of the *TRAPPC10* variants, including the possibility that the protein C-terminus plays a specific though as yet unknown post-Golgi secretory pathway role. Interestingly, in previous studies the yeast homologue of TRAPPC10 (Trs130) was rendered conditionally lethal by short truncations in the C-terminal portion of the protein, indicative of a critical functional role of this region [258]. The observed trafficking defects herein may be explained by the loss of GEF activity of TRAPP II towards Rab1 [266], Rab11 [264] and/or Rab18 [267], which alongside other Rab GTPases are key regulators of cellular trafficking by recruiting specific effector proteins to membranes, including lipid droplet homeostasis. Rab18 has specifically been implicated in secretory pathway regulation [326], and *RAB18* variants are associated with a neurodevelopmental disorder (OMIM #614222 [327]) phenotypically overlapping *TRAPPC10*-related disorder. Although the specific molecular role of TRAPPC10 in neurological development and function remains unclear, it is tempting to speculate that the predominantly neurological phenotype seen in individuals affected by *TRAPPC10*-related disorder may be due to the relatively high rates of membrane trafficking shown to be required in the brain for processes such as synapse remodelling [328, 329].

An important outcome of our findings was that absence of TRAPPC10 is also associated with a concomitant absence of TRAPPC9. Additionally, although other TRAPP III and core specific components were detected, levels of core TRAPPC2L were reduced suggesting an absence of TRAPP II complex, with no effect on TRAPP III complex. These findings suggest that TRAPPC10 is crucial to TRAPP II complex stability and function. Consistent with our findings, careful examination of the data in Li *et al* (2017), though not commented on, reveals a decrease in the levels of TRAPPC9 protein in *TRAPPC10*$^{-/-}$ HEK293 cells (see Figure 6D in Ref 4) [267]. However, our studies showed TRAPPC10 was still detectable in LCLs lacking

TRAPPC9 (p.(Arg475*)), suggesting that TRAPPC9 may only incorporate into TRAPP II in the presence of TRAPPC10, while the reciprocal is not true. Our overexpression studies of mutant TRAPPC10 provide further insight into this, and in particular identify the expression of full length TRAPPC10 to be of importance for the detection of TRAPPC9 within the TRAPP II complex. Collectively, our results suggest that the absence of the TRAPPC10 C-terminus results in TRAPPC10 protein degradation, and consequent loss of TRAPPC9 (through protein degradation or some other mechanism), leading to the absence of TRAPP II.

The concomitant loss of TRAPPC9 associated with pathogenic *TRAPPC10* variants is reflected in the close phenotypical parity of human *TRAPPC9-* (OMIM #613192) and *TRAPPC10*-related conditions (Table 3.1) and with both knockout mouse models. Individuals with *TRAPPC9*-related disorder have postnatal-onset microcephaly [296, 299]. While the families reported here did not have any recorded birth OFC measurements, there were no reports of affected children having noticeably small head size at birth suggesting microcephaly associated with *TRAPPC10*-related disorder may be postnatal. Head growth trajectory was only available for one child, with two measurements available (2.5 and 4.1 years), that indicate microcephaly is present from early childhood and is not progressive. However it remains unclear whether the brain growth was initially normal and then slowed, or whether microcephaly was present from/before birth. Identification of further affected families will help clarify the nature of the microcephaly in *TRAPPC10*-related disorder, as with *TRAPPC9*-related disorder where this was also unclear initially [293].

Biallelic TRAPPC10 loss of function appears to have a more severe phenotypic impact in mice and humans, compared with biallelic TRAPPC9 loss. In human *TRAPPC10-* related disorder, microcephaly appears to be a relatively consistent feature, (8/10 individuals), as compared to ~60% of TRAPPC9 patients [294, 330]. Similarly in mouse models, while the age of onset and progression of microcephaly is unclear, *Trappc10$^{-/-}$* mice display more extensive brain size reduction and loss of white matter structures than *Trappc9$^{-/-}$* mice. These findings may be explained by the concomitant loss of TRAPPC9 observed with TRAPPC10 absence, likely acting as a double knockout. Obesity is reported in ~50% of individuals with *TRAPPC9*-related disorder and is the only clinical feature not currently observed in individuals with *TRAPPC10-* related disorder. It is important to note that these differences are based on a relatively

small sample size, a limitation of this study. Nevertheless, the *Trappc10⁻ᐟ⁻* mice share similar phenotypic outcomes compared to individuals with biallelic *TRAPPC10* variants.

Neuroanatomical studies of *Trappc9⁻ᐟ⁻* mice showed overlapping features with *Trappc10⁻ᐟ⁻* mice including reduced brain size predominantly of white matter structures (Appendix 2, Figure A2.4) [331], although findings of enlarged striatal size were not identified in *Trappc10⁻ᐟ⁻* mice. TRAPPC9 plays a role in the NF-κB signalling pathway through its interaction with NIK and IKKβ, which both regulate the NF-κB pathway [275, 332]. Given the associated loss of TRAPPC9 observed with both the *TRAPPC10* variants and TRAPPC10⁻ᐟ⁻ cells, it is plausible that TRAPPC10 loss of function may also impact NF-κB signalling pathways through disruption of TRAPPC9, although the involvement of TRAPPC9 in this pathway in brain has been recently questioned [331]. Taken together, we define the genetic, clinical and molecular basis of a novel microcephalic neurodevelopmental disorder associated with biallelic *TRAPPC10* variants in both humans and mice. Our data provides a molecular rationale for the phenotypic overlap between *TRAPPC10-* and *TRAPPC9-*related TRAPPopathy disorders, both involving disruption of TRAPP II-mediated post-Golgi trafficking processes.

The compelling evidence presented here of variants within *TRAPPC10* as causative of a microcephalic neurodevelopmental TRAPPopathy disorder will help to facilitate other individuals with *TRAPPC10*-related disorder to achieve an accurate molecular diagnosis, ending the diagnostic odyssey for affected individuals and their families and enabling the provision of informed genetic counselling. *TRAPPC10*-related disorder should be considered in the differential diagnosis for patients with microcephaly and cognitive impairment.

3.6 Conclusions and future work

This chapter details the clinical, genetic and functional characterisation of *TRAPPC10*-related disorder and identifies a striking similarity in phenotype with a mouse model of the disorder. The approach used in this study to combine whole-genome SNP mapping in a number of affected individuals with WES in a single affected individual in a large family provided a powerful and cost-effective method to identify that a homozygous variant in *TRAPPC10* is associated with a severe neurodevelopmental disorder. This enabled the subsequent identification of two further families with the condition (one previously published), to consolidate variants in this gene as a cause of an autosomal recessive form of microcephalic neurodevelopmental disorder. Confirmation of a molecular diagnosis is not only hugely beneficial to the families participating in this study to understand the cause of their condition and receive appropriate genetic counselling, but also to other families in the local community and patients worldwide who may be diagnosed with *TRAPPC10*-related disorder now this condition has been described. Future studies should attempt to identify additional affected individuals and pathogenic *TRAPPC10* variants to further characterise the full clinical and genetic disease spectrum associated with *TRAPPC10*-related disorder and help to aid the development of specific clinical management guidelines for this condition. Further studies aiming to identify the underlying molecular cell interactions and signalling pathways involved in development of *TRAPPC10*-related disorder should be conducted and attempt to identify any possible therapeutic targets for potential treatment of this condition.

Studies were undertaken to identify whether variants in other TRAPP complex subunits are associated with human disease, which led to the discovery of three further TRAPP complex subunits (*TRAPPC1*, *TRAPPC5* and *TRAPPC8*) as candidate TRAPPopathy disorder genes, and candidate variants which with further work would consolidate two TRAPP complex subunits (*TRAPPC3* and *TRAPPC14*), each previously described in only a single family, as causative of neurodevelopmental TRAPPopathies. Additional clinical, genetic and functional studies are required to confirm the findings in these families associated with TRAPP complex (1, 3, 5, 8 and 14) variants. This includes detailed studies to compile more complete clinical and family history datasets for each family, alongside more thorough genetic studies to evaluate likely pathogenicity of the TRAPP molecule variants and exclude other

genetic variants. Additionally cell and molecular studies such as trafficking assays (Appendix 2, A2.1.1) will be very helpful for defining any functional outcomes and confirming the likely pathogenicity of each variant. Studies of mouse knockout models of these TRAPP complex molecules will also provide further evidence of pathogenicity similar to the *Trappc10*^{-/-} mouse studies described here and we hope to continue with these collaborations to conduct these important studies.

4 CLINICAL, GENETIC AND MOLECULAR DELINEATION OF *KPTN*-RELATED DISORDER IN HUMANS AND MICE

4.1 Summary

KPTN-related disorder (also known as macrocephaly, autistic features, seizures, developmental delay; MASD syndrome) is an autosomal recessive disorder associated with pathogenic biallelic germline variants in the *KPTN* gene encoding kaptin, a key molecular component of the mTOR regulatory complex KICSTOR, first identified in the Amish community by our research group in 2014. This study identifies 21 new patients with this condition, revisits nine of the originally described Amish patients, and reviews data for six patients in previously published studies. To gain further insights into the pathomechanistic basis of this condition, mouse and human iPSC KPTN loss-of-function models were investigated. *Kptn*$^{-/-}$ mice display many of the key phenotypical features seen in the human disorder, including megalencephaly and behavioural abnormalities, as well as specific cognitive deficits not originally recognised in humans. This led to re-evaluation of these aspects in patients, which revealed concordant specificity of cognitive deficits with relative sparing of hippocampal-independent memory, postnatal progressive brain overgrowth, and a previously unrecognised dosage-sensitivity to loss of KPTN protein, affecting head circumference of heterozygous carriers. Molecular and structural analysis of mouse brain tissue revealed pathological changes contributing to the mouse phenotypes, including differences in brain size, shape, and cell numbers primarily due to abnormal postnatal brain development. Both the mouse and differentiated iPSC models of the disorder display transcriptional and biochemical evidence for altered mTOR pathway signalling, supporting the role for KPTN in regulating mTORC1.

Together these studies provide a more detailed characterisation of the genetic basis and phenotypical spectrum of *KPTN*-related disorder, and enable the condition to be positioned in the broader group of mTORC1 related disorders "mTORopathies", characterised by abnormalities in brain structure, cognitive function, and seizures. Importantly, as increased mTOR signalling downstream of KPTN is rapamycin sensitive, this highlights possible therapeutic avenues for the condition utilising currently available mTOR inhibitors.

4.2 Introduction

4.2.1 *KPTN*-related disorder

The Windows of Hope project and their collaborators previously described inherited biallelic variants in the *KPTN* gene as the cause of a clinically recognisable novel autosomal recessive macrocephalic neurodevelopmental disorder in nine affected individuals from four nuclear families within an extended interrelated family of Ohio Amish background (Figure 4.1). These studies identified that *KPTN*-related disorder (mental retardation, autosomal recessive 41; OMIM #615637) is characterised by cardinal clinical features of global developmental delay, intellectual disability, macrocephaly and behavioural problems including, autistic features, anxiety, stereotypies, hyperactivity and repetitive speech [333]. Additional features include seizures (generalised tonic-clonic, absence and complex partial seizures), recurrent infections, hepatomegaly, splenomegaly, childhood hypotonia, and dysmorphic features. The condition was associated with two distinct loss of function (LOF) founder variants in *KPTN* (NM_007059.4 Chr19:g.47479874G>T, c.776C>A; p.(Ser259*) and Chr19:g.47479918G>GACCGACCACATCTGCAGA, c.714_731dup; p.(Met241_Gln246dup) [hg38]), with affected individuals being either homozygous or compound heterozygous for each variant (Figure 4.1). Subsequently, four further clinical studies have been published describing six individuals with phenotypic features overlapping those identified in the Amish [334-337].
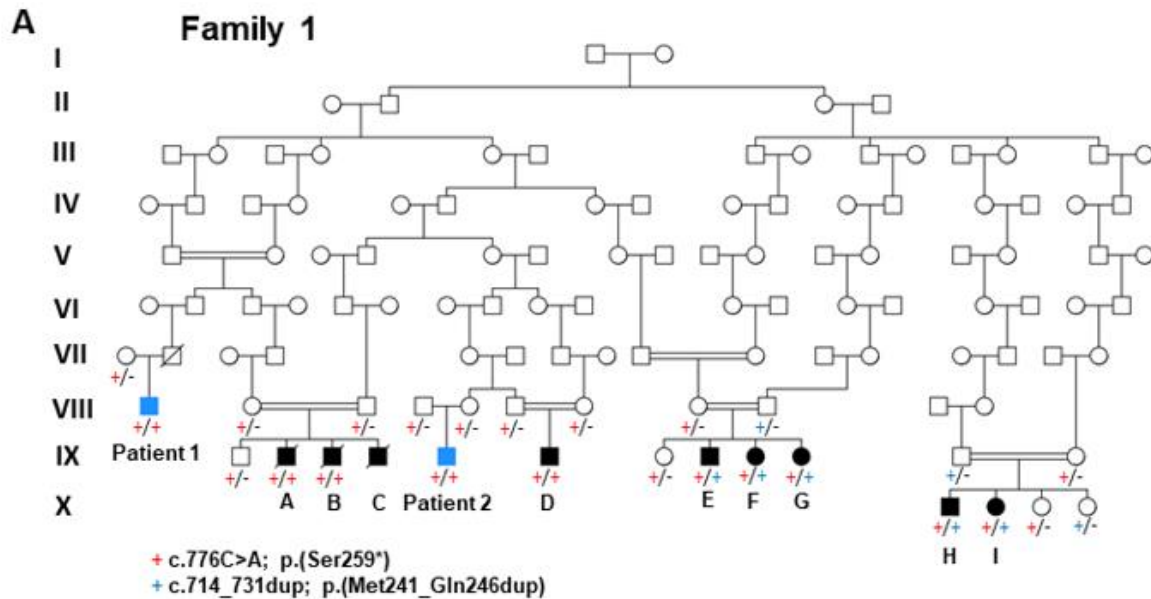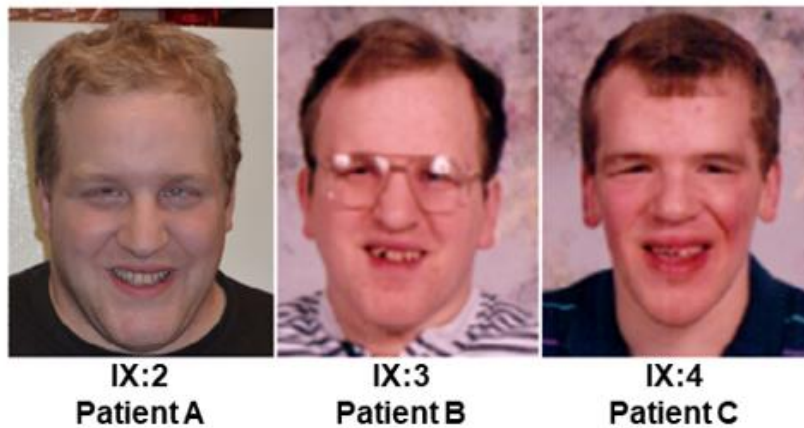
Figure 4.1: Pedigree and clinical photographs of affected Amish individuals from Family 1

(**A**) Pedigree of Family 1, an extended inter-related Amish family with 11 individuals affected with *KPTN-*related disorder (shaded symbols). Individuals in black were previously published by our group in 2014 [333] and individuals coloured blue are two newly identified individuals. Genotype segregation is shown under each individual '+', indicates the presence of a *KPTN* variant, '-', wild type allele. Ohio Amish founder variants are identified by red '+' c.776C>A; p.(Ser259*) and blue '+' c.714_731dup; p.(Met241_Gln246dup). (**B**) Photographs of three brothers (IX:2-4) with *KPTN*-related disorder, homozygous for the c.776C>A; p.(Ser259*) variant and first described in Baple *et al* (2014), showing facial features of macrocephaly with frontal bossing and a high frontal hairline, downslanting small palpebral fissures (PF), long face with a prominent chin, broad nasal tip, and hooded eye lids. Adapted from Baple *et al* [333].

Pajusalu *et al.* (2015) described two adult Estonian siblings from a geographically isolated region with homozygous inheritance of a frameshift *KPTN* variant ((NM_007059.4: c.665dupA; p.(Glu222fs*)), associated with a phenotype similar to that previously described [334]. Thiffault *et al.* (2018) reported a 9-year-old male with focal epilepsy, static encephalopathy, autism and macrocephaly, with additional features of hepatosplenomegaly, hypoglycaemic episodes and severe intractable epilepsy resulting in the boy's death from status epilepticus aged nine [338]. Genetic analysis identified compound heterozygosity of the p.(Met241_Gln246dup) variant described in the Amish, alongside a splice variant (NM_007059.4: c.394+1G>A; p.?) previously reported in a patient with multiple congenital anomalies, although no further clinical details were available [339]. Two further publications reported three individuals with the same homozygous *KPTN* variant (NM_007059.4; c.597_598dup; p.(Ser200Ilefs*55)) associated with postnatal macrocephaly, neurodevelopmental delay and intellectual disability [336, 337]. One of these individuals displayed an additional feature of precocious puberty with hyperprolactinaemia and two had delayed anterior fontanelle closure. A total of four *KPTN* gene variants were identified in these subsequent studies, including the p.(Met241_Gln246dup) variant originally identified in the Amish.

4.2.2 KPTN expression and function

The *KPTN* gene located on chromosome 19q13.32 entails 12 exons encoding kaptin, a 436 amino acid polypeptide [340] containing an integrin alpha N-terminal domain [341]. The molecule has homologues in most mammalian, bird and fish species as diverse as *Ciona intestinalis* [342]. Bearer *et al.* (1992) first identified the KPTN protein (2E4) in rats, whilst investigating the role of actin polymerisation in haemostasis [343]. KPTN was found to be highly expressed in human infant and fetal brain [344]. Figure 4.2 (GTEx; https://gtexportal.org/home/gene/KPTN) shows that tissue expression of KPTN is high within all brain regions, with highest expression within the cerebellum, and also within the pituitary, testis, prostate and cervix.
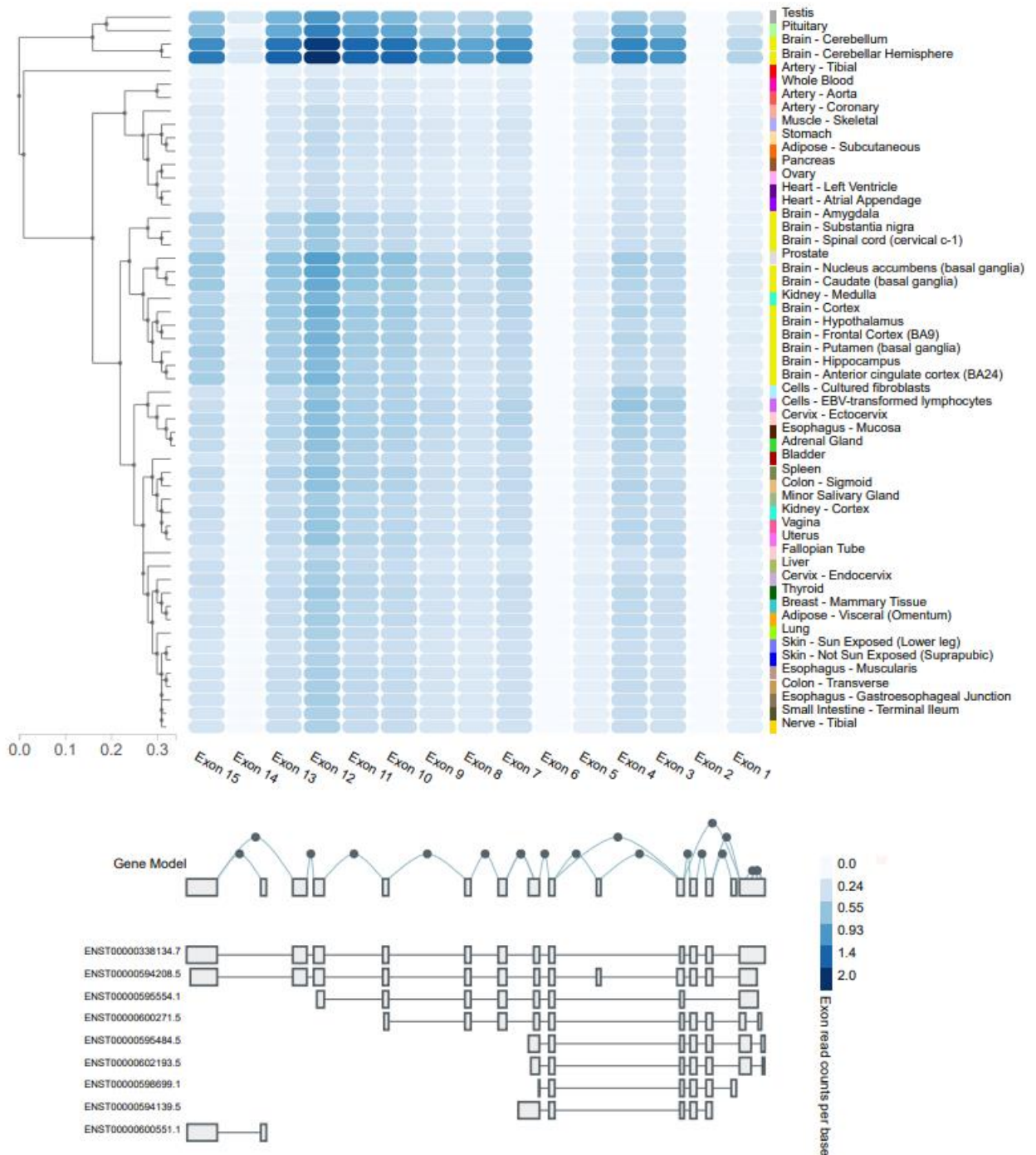
Figure 4.2: Expression levels of KPTN within human tissues

Expression plot showing expression levels of KPTN in each exon. Dark blue indicates high expression levels (2.0 exon read counts per base) and white indicates absent expression (0.0 exon read counts per base). Reproduced from the Genotype-Tissue Expression (GTEx) Project (supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS) obtained from:https://gtexportal.org/home/gene/KPTN, accessed on 25/05/2022.

Actin cytoskeleton

KPTN was found to regulate the growth of filamentous (F)-actin, by binding to its fast-growing ends at sites of active actin polymerisation, including neuronal growth cones [340, 344]. The actin cytoskeleton plays a crucial role in neuronal cell migration and morphology during neurodevelopment, including the generation of growth cones and dendritic spines [345-347]. Genetic variants in other proteins involved in regulating actin filament growth and dendritic spine formation entail a well-defined cause of neurodevelopmental disorders [348-350].

4.2.3 mTOR signalling, KPTN and the KICSTOR complex

mTOR (mechanistic target of rapamycin), is an atypical serine/threonine kinase within the phosphoinositide 3-kinase (PI3K) family, that forms the catalytic subunit of two distinct protein complexes, known as mTOR Complex 1 (mTORC1) and 2 (mTORC2) [31, 146, 351, 352]. mTOR signalling regulates many important and distinct cellular processes including protein synthesis, cell growth, and nutrient uptake. Moreover, mTOR signalling is central to the regulation of long lasting synaptic plasticity, which is reliant on protein synthesis and is critical for the formation and persistence of memories [353]. Due to the pivotal roles of mTOR signalling in many developmental and homeostatic processes, this pathway has been implicated in the aging process [354] and the pathogenesis of numerous diseases such as cancer, diabetes, epilepsy, as well as neurodegenerative disorders including Alzheimer's, Parkinson's, Huntington's diseases, and neuropsychiatric disorders including depression, schizophrenia and autism [31, 146, 351, 352]. *De novo* and inherited pathogenic variants in numerous mTOR pathway genes (Figure 4.3) are associated with the "mTORopathies", a group of disorders in which disruption of the mTOR signalling pathway results in conditions with common phenotypes including epilepsy, neurodevelopmental delay and megalencephaly [146, 355-359]. A growing body of literature is defining the intricate role of the mTOR pathway in neural stem cell proliferation and neuronal differentiation [360-363], providing insights into the pathomechanism of mTOR-related brain disorders such as tuberous sclerosis, focal cortical dysplasia, hemimegalencephaly, and Smith-Kingsmore syndrome [364-366] (Table 4.1).
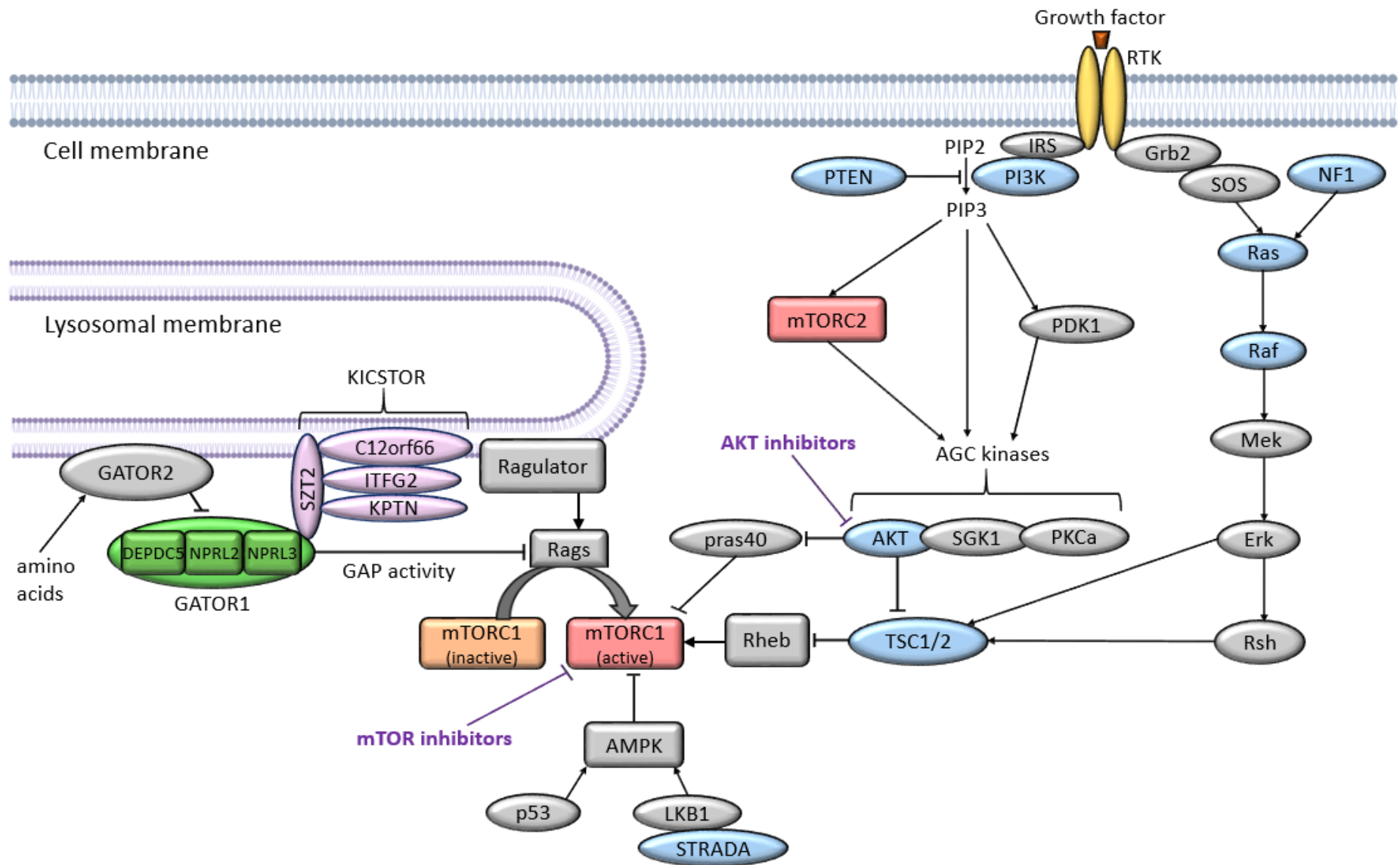
Figure 4.3: Schematic diagram of the proposed pathways involved in regulation of mTOR activation

mTOR is an atypical serine/threonine protein kinase within the phosphoinositide 3-kinase (PI3K) family that forms two complexes mTORC1 and mTORC2. The KICSTOR complex is composed of four subunits: KPTN, ITFG2, C12orf66 and SZT2 (purple proteins) [367], and interacts with and recruits GATOR1 (composed of subunits DEPDC2, NRPL2/3, green proteins) to the lysosome membrane where it inhibits Rag activation of mTORC1 through GTPase activating protein (GAP) activity [368]. Pathogenic variants in genes encoding the proteins in blue have been associated with neurodevelopmental disorders: STRADA; polyhydramnios, megalencephaly, and symptomatic epilepsy (PMSE), TSC1/2; tuberose sclerosis types 1 and 2, AKT1-3; Cowden syndrome 6, Hypoinsulinaemic hypoglycaemia with hemihypertrophy, Megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome 2, PTEN; Cowden syndrome 1, Lhermitte-Duclos syndrome, Macrocephaly-autism syndrome, PI3K; Megalencephaly-capillary malformation-polymicrogyria syndrome (somatic), Cowden syndrome 5, NF1; neurofibromatosis type 1, Ras/Raf; Noonan syndrome. mTOR inhibitors act directly on mTORC1 to inhibit activation and AKT inhibitors are another group of potential therapeutic treatment. Adapted with permission from Wolfson *et al* (licence granted by Springer Nature, Appendix 7).

GENE	KPTN	SZT2	DEPDC5	NPRL2	NPRL3	STRADA	TSC1/TSC2	PTEN	mTOR
Disorder	*KPTN*-related Syndrome	Early infantile epileptic encephalopathy 18	Familial focal epilepsy with variable foci 1	Familial focal epilepsy with variable foci 2	Familial focal epilepsy with variable foci 3	Polyhydramnios, megalencephaly and symptomatic epilepsy	Tuberous sclerosis type 1/2	Cowden/ Macrocephaly-autism syndrome	Smith-Kingsmore Syndrome
OMIM #	615637	615476	604364	617116	617118	611087	191100	158350	616638
Inheritance	AR	AR	AD	AD	AD	AR	AD	AD	AD
Neurological features									
Developmental delay	Mild-severe	Mod-Severe	Variable	Variable	Variable	Severe	Variable	Mild-mod	Mild-severe
Intellectual disability	Mild-severe	Mod-Severe	Variable	Variable	Variable	Severe	Variable	Mild-mod	Mild-severe
Macrocephaly	✓	✓	Hemimegalencephaly			✓	✓	✓	✓
Hypotonia	✓	✓				✓			✓
Seizures	GTCS, absence, variable severity	GTCS, absence, focal, intractable, early onset	Focal/multifocal, secondary generalisation, nocturnal seizures	Focal/multifocal, secondary generalisation, nocturnal seizures	Focal/multifocal, secondary generalisation, nocturnal seizures	GTCS, early onset, refractory	GTCS, infantile spasms	GTCS	GTCS, intractable
Neuroimaging	Mild ventriculomegaly	Persistent cavus septum pellucidum, short thick corpus callosum, enlarged ventricles	Focal cortical dysplasia	Focal cortical dysplasia	Focal cortical dysplasia	Ventriculomegaly	Hamartomatous lesions, cortical tubers, intracranial calcification	Cerebellar gangliocytoma, Lhermitte-Duclos disease	Ventriculomegaly, dysgenesis of the corpus callosum, mild dysmyelination
EEG	Generalised slowing, bifrontal spike and slow wave	Slowed background activity, abnormal spike waves	Sharp waves, spikes or spike-and-waves, focal or multifocal, focal slow waves	Sharp waves, spike-and-waves, focal or multifocal, focal slow waves	Sharp waves, spike-and-waves, focal or multifocal, focal slow waves	Multifocal spike and slow wave			
Behavioural features									
Autistic features	✓		✓		✓		✓	✓	✓
Anxiety	✓		✓						
Hyperactivity	✓		✓				✓		✓
Facial features									
Frontal bossing	✓	✓				✓		✓	✓
Prominent chin	✓					✓			
Hypertelorism	✓	✓				✓			✓
Downslanting palpebral fissures	✓	✓							✓

	C1	C2	C3	C4	C5	C6	C7	C8	C9
High arched palate	✓	✓						✓	
Other									
Hepatomegaly	✓							✓	
Splenomegaly	✓							✓	
Recurrent infections	✓ IgA deficiency							✓	✓ IgA def
Strabismus	✓					✓			✓
Neoplasia			✓		✓		✓	✓	
SUDEP	✓		✓	✓	✓	✓	✓		
Treatment									
mTOR inhibitors	N/P	N/P	N/P	N/P	N/P	Rapamycin – reduced seizure frequency, improved receptive language and alertness [369, 370]	Rapamycin, everolimus – reduced seizure frequency and tumor volume [371]	Rapamycin – improvement in hamartomas and vascular anomalies [372-374]	N/P

Table 4.1: Clinical features of mTORopathies associated with hyperactivation of mTORC1

Shaded columns represent different mTOR pathway complexes, purple: KICSTOR, green: GATOR1, yellow: downstream mTOR regulators, red: mTOR protein.

Abbreviations: PMSE: polyhydramnios, megalencephaly and symptomatic epilepsy, AR: autosomal recessive, AD: autosomal dominant, GTCS: generalised tonic-clonic seizures, EEG: electroencephalogram, SUDEP: sudden unexplained death in epilepsy, IgA: Immunoglobulin A, N/P: none published.

In 2017, Wolfson *et al* determined that KPTN is a core component of the KICSTOR complex alongside three other subunits; ITFG2 (integrin alpha FG-GAP repeat containing 2), C12orf66 (chromosome 12 open reading frame 66), and SZT2 (seizure threshold 2), which is involved in upstream negative regulation of mTORC1 (the mechanistic target of rapamycin complex 1) activation [367] (Figure 4.3). Within the KICSTOR complex, KPTN and ITFG2 form a heterodimer that interacts with C12orf66 only in the presence of SZT2, suggesting a linking role for SZT2 [367]. Notably, biallelic pathogenic variants in *SZT2* cause a neurodevelopmental disorder characterised by severe early-onset epileptic encephalopathy, global developmental delay, structural brain abnormalities such as thick and short corpus callosum, and macrocephaly [319, 320]. In mice, Szt2 loss-of-function mutants display epileptogenesis and increased mTORC1 signalling in the brain [318, 338]. In cell lines, KICSTOR is required for the inhibition of mTORC1 signalling in response to amino acid or glucose deprivation [352, 367]. mTORC1 is regulated by the GATOR1 (Gap Activity TOward Rags 1) protein complex through GTPase activating protein (GAP) activity that inhibits Rag activation of mTORC1 in response to amino acid substrate levels [368, 375]. The KICSTOR complex is important for the recruitment of GATOR1 to the lysosome in an amino acid-independent manner (Figure 4.3) and importantly loss of KICSTOR components has been shown to inhibit GATOR1 recruitment and increase mTORC1 signalling in murine neurons, hepatocytes and cardiac myocytes [367].

mTORC1 is a pivotal regulator of cerebral cortical development through promoting neuronal differentiation, neurite elongation, branching and synapse formation in response to multiple intra-and extra-cellular signals including growth factors, amino acids, glucose, hypoxia, neurotransmitters/modulators and hormones [37]. Several neurodevelopmental disorders have been associated with hyperactivation of mTORC1 (Table 4.1, Figure 4.3), including the autosomal recessive conditions: early infantile epileptic encephalopathy type 18 [356, 376] and polyhydramnios, megalencephaly, and symptomatic epilepsy (PMSE) [370, 377] caused by biallelic variants in *SZT2* and *STRADA* respectively (Table 4.1). Several autosomal dominant mTORopathies have been identified including familial focal epilepsy with variable foci 1, 2 and 3 associated with variants in GATOR1 subunits *DEPDC5, NPRL2* and *NPRL3* respectively [378, 379], tuberose sclerosis types 1 and 2 (*TSC1/2*) [140, 351] and several others [351, 365, 378] with overlapping features including megalencephaly, intellectual disability

and seizures (Table 4.1). In addition, several somatic overgrowth disorders have been associated with key proteins within the PI3K/AKT/mTOR pathway, including megalencephaly-capillary malformation-polymicrogyria syndrome associated with somatic *PIK3CA* variants [142].

This study expands the genotypic and phenotypic spectrum in 36 individuals affected with *KPTN*-related disorder. As this condition has not been previously modelled in mice or cells, we sought to assess the functional consequences of homozygous *Kptn* LoF in mice, as well as human iPSC models of the condition. Additionally, we propose clinical management guidelines based on the results of this work and explore potential future therapy options through mTOR inhibition.

4.3 Methods

4.3.1 Clinical and genomic studies

Clinical and genomic studies were carried out as described in sections 2.2 and 2.4. Appendix 4, section A4.1.1 describes the identification of affected individuals. Detailed psychometric testing was performed for a subset of Amish patients and age-matched controls from the same community (section 2.2.2.1), to provide community appropriate control data. Psychometric testing was undertaken by Dr Emma Baple with support from Dr James Tonks (Consultant Paediatric Neuropsychologist and Consultant Clinical Psychologist, University of Exeter Medical School). I analysed the psychometric test results with support from Dr Tonks. Genomic studies (WES and dideoxy sequencing) identifying *KPTN* variants in new non-Amish affected individuals was undertaken diagnostically or by collaborating groups. SNP genotyping was undertaken as previously described (section 2.4.6). Mutational age analyses were undertaken with support from Dr James Fasham (University of Exeter), see Appendix 4, section A4.1.1.

4.3.2 Mouse, molecular and cell studies

Mouse and functional studies were designed by Dr Sebastian Gerety, Dr Darren Logan and Dr Chris Lelliott (Wellcome Trust Sanger Institute, UK), Professor Peter Crino and Dr Philip Iffland (University of Maryland, USA) and Dr Binnaz Yelcin (Inserm, France) after discussion of the human phenotype with myself, Dr Baple and Professor Crosby.

The analysis was primarily undertaken by Dr Maria Levitin, alongside other collaborating groups.

Methodology for the mouse, molecular and cell studies are detailed in Appendix 4 A4.1.2 and A4.1.3.

4.4 Results

4.4.1 *Kptn*^{-/-} mice recapitulate *KPTN*-related disorder patient phenotypes

These studies used an engineered loss-of-function mouse model (Kptn$^{tm1a(EUCOMM)Wtsi}$) generated by the Wellcome Sanger Institute's Mouse Genetics Project [380] (Appendix 4 A4.2.1 Generation and characterisation of a *Kptn*^{-/-} mouse model). The resulting allele closely resembles the truncating allele (p.(Ser259*)) observed in the original *KPTN*-related disorder publication [333], resulting in premature termination in Exon 8.

4.4.1.1 *Kptn*^{-/-} mice have increased locomotor activity and anxiety-like phenotypes

To investigate whether the mouse model recapitulates any of the neurocognitive features of the human *KPTN*-related disorder, we tested cohorts of *Kptn*^{-/-} mice through a series of functional tests. Locomotor capabilities, including distance covered and time spent moving, were assessed in *Kptn*^{-/-} mice compared to wildtype controls, using video assisted observation and scoring in an open field test [381]. *Kptn*^{-/-} mice travelled a significantly greater distance than *Kptn*^{+/+} controls, accounted for by the increase in time spent moving ($p < 0.05$, Figure 4.4A,B). These findings are consistent with our observations of hyperactivity in *KPTN*-related disorder patients [333, 334, 336-338]. Other activity parameters including velocity travelled were not significantly different between genotypes (Appendix 4, A4.2.1).
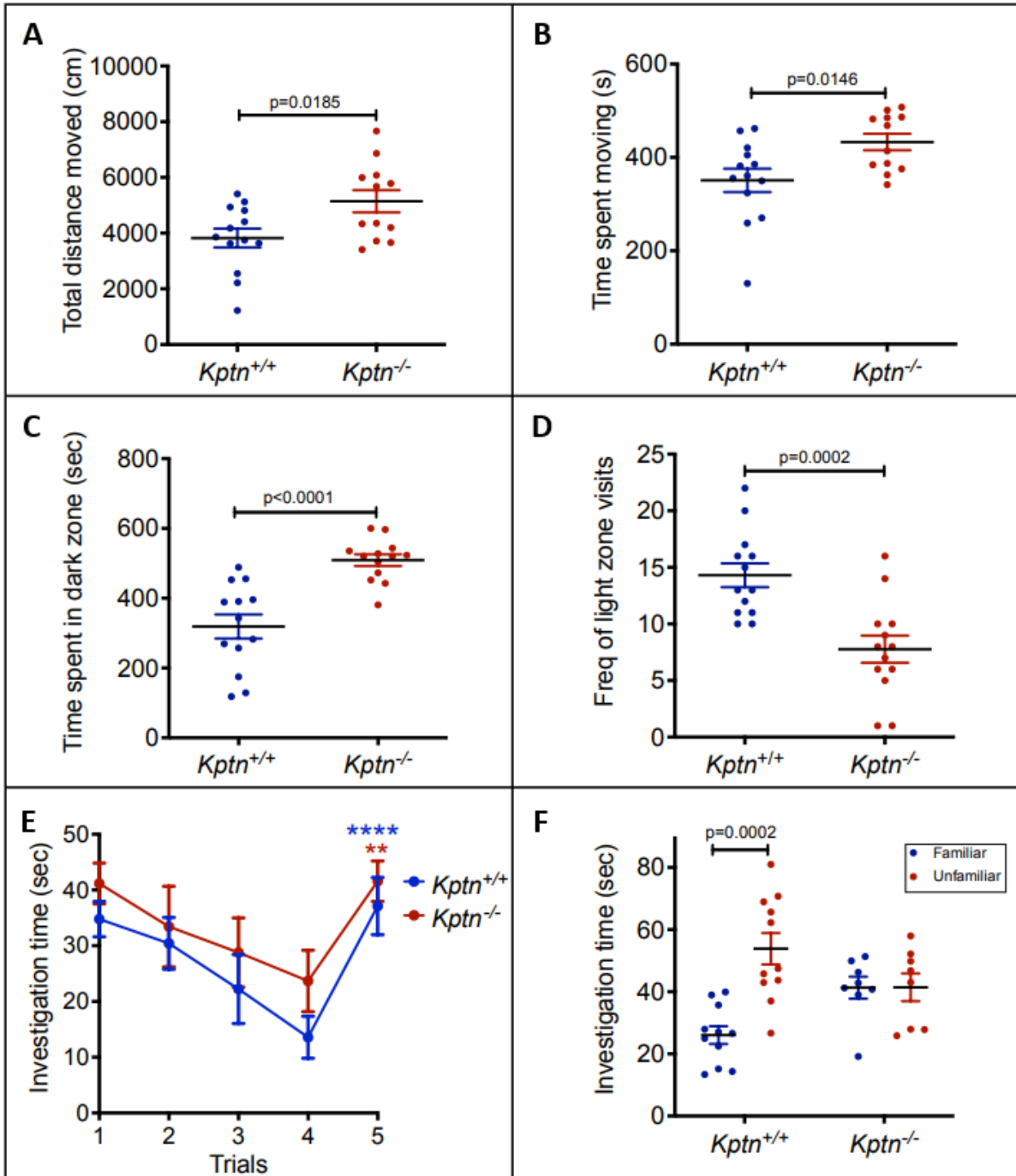
Figure 4.4: *Kptn⁻/⁻* mice display behavioural, anxiety and memory deficits

Kptn⁻/⁻ mice display increased locomotor activity (**A,B**), anxiety-like behaviour (**C,D**), and reduced memory retention (**E,F**). (**A,B**) The distance covered in an open field and time spent moving. (**A**) *Kptn⁻/⁻* mice (red; n=12) covered significantly more distance (p=0.0185, t=2.534 df=23, two-tailed Student's t-test) than wildtype controls (*Kptn⁺/⁺*, blue; n=13) and (**B**) spent more time moving (p=0.0146, t=2.64 df=23, two-tailed Student's t-test). (**C**) Kptn⁻/⁻mice (red; n=14) spend significantly longer time in the dark zone (p<0.0001, t=4.946 df=24, two-tailed Student's t-test) of a light/dark box than wildtype controls (*Kptn⁺/⁺*, blue; n=13), and (**D**) have reduced frequency of visits to the light zone (p=0.0002, t=4.326

df=25, two-tailed Student's t-test). (**E**) Social recognition (Day 1). Both controls (*Kptn*$^{+/+}$, n=11) and *Kptn* mutant mice (*Kptn*$^{-/-}$, n=8)) recognise a stimulus animal repeatedly presented to them over the course of four trials, as shown by a decline in the investigation time over trials 1-4. Both mutant and wildtype mice display an increase in the investigation time on trial 5 when presented with a novel stimulus animal (two-way ANOVA, interaction TRIAL x GENOTYPE $F_{(4, 68)}$ = 0.3852, P=0.8185; TRIAL $F_{(4, 68)}$ = 16.26, P<0.0001; post-hoc *Kptn*$^{+/+}$ p<0.0001****, *Kptn*$^{-/-}$ p=0.0023**). (**F**) Social discrimination (Day 2). Wildtype *Kptn*$^{+/+}$ controls (n=11), but not *Kptn*$^{-/-}$ mice (n=8) show a significant increase in time spent investigating an unfamiliar stimulus versus the familiar stimulus mouse from Day 1, suggesting that *Kptn*$^{-/-}$ mice do not retain the memory of a familiar animal at 24h (*Kptn*$^{+/+}$: p=0.0002, t= 4.79 df=2; *Kptn*$^{-/-}$: p= 0.985, t=0.0185 df=14, two-tailed multiple t-test with multiple comparison correction). Values are plotted as mean ± SEM. Data courtesy of Dr Sebastian Gerety [382].

To assess anxiety in *Kptn*$^{-/-}$ mice, we performed a light/dark box test [383, 384]. *Kptn*$^{-/-}$ mice spent significantly more time in the dark zone (59.6% increase, p=4.8X10^{-5}), and had reduced frequency of visits to the light zone (45.6% reduction, p=0.0006), (Figure 4.4C,D). These results indicate a strong anxiety-like phenotype in the *Kptn*$^{-/-}$ mouse model when compared to wildtype controls, concordant with the anxiety observed in patients with *KPTN*-related disorder.

4.4.1.2 *Kptn*$^{-/-}$ mice have cognitive deficits

To address whether *Kptn*$^{-/-}$ mice display intellectual deficits concordant with the human disorder, learning and memory capacity was evaluated using an ethologically relevant social recognition (SOR) assay, which assesses olfactory-mediated hippocampus-dependent memory [385-387]. *Kptn*$^{-/-}$ mice had no detectable deficit in social interaction compared with their wildtype counterparts on day 1, as measured by the time spent investigating the novel conspecific (stimulus A) on trial 1 (Figure 4.4E). With repeated exposure over four trials, both wildtype and *Kptn*$^{-/-}$ mice habituated to the stimulus animal (stimulus A), indicating a functional olfactory system and both increased their investigation time of a novel mouse (stimulus B, trial 5), indicating successful recognition (Figure 4.4E). On day 2, when given a choice between investigating an unfamiliar mouse and a familiar one (stimulus A), wildtype mice spent significantly more time investigating the unfamiliar mouse (p=0.0002) than *Kptn*$^{-/-}$ mice that did not show a preference for either stimulus (non-significant p=0.985, Figure 4.4F), indicating a 24 hour memory impairment. The *Kptn*$^{-/-}$ mice had a significantly reduced preference index for the unfamiliar stimulus (53.9%, p=0.0136) compared to *Kptn*$^{+/+}$ mice (67.5%). These results of the SOR assay indicate a short-term hippocampal dependent memory deficit in *Kptn*$^{-/-}$ mice.

Further assessment of hippocampus-dependent spatial memory in *Kptn*$^{-/-}$ mice was conducted using the Barnes maze [388], which found significant impairment in long term hippocampus-dependent spatial memory compared with wildtype mice; results are detailed in Appendix 4 (A4.2.1 *Cognitive deficits in Kptn*$^{-/-}$ *mice*). To test whether *Kptn*$^{-/-}$ mice also display impairments in tasks less reliant on hippocampal function, pairwise discrimination (PD) was tested, a visual operant conditioning task using a touchscreen platform [389], with no significant difference in the performance of *Kptn*$^{-/-}$ compared with wildtype mice (see Appendix 4, A4.2.1, Figure A4.2). Together, the

results of three independent memory tests suggest that $Kptn^{-/-}$ mice experience selective deficits, with at least some hippocampal-independent memory sparing.

4.4.1.3 *KPTN* deficiency is associated with severe and progressive macrocephaly in mice

Features of macrocephaly and frontal bossing identified in *KPTN*-related disorder patients were assessed in $Kptn^{-/-}$ mice. Microcomputed tomography (µCT) [390], magnetic resonance imaging (MRI) and histological morphometrics were used to quantify changes in skull and brain volume, shape, and cellular distribution [390]. Assessment of the skulls of adult $Kptn^{-/-}$ mice identified an increase in height, length, and width of the brain cavity (Figure 4.5A-D) with further details in Appendix 4 (A4.2.1 *Neuroimaging identifies changes in skull and brain volume in Kptn^{-/-} mice*). An increase in the dorsal curvature of the frontal and parietal bones is consistent with the changes observed in frontal bossing in humans, where the frontal bone can expand to accommodate the mechanical pressure of pathological brain overgrowth [391, 392]. There was also a significant increase in mean intracranial volume in $Kptn^{-/-}$ mice when compared to wildtype controls (Figure 4.5E). Histomorphological analysis of neuroanatomical structures are detailed in Appendix 4 (A4.2.1 *Histomorphological analysis of neuroanatomical structures in Kptn^{-/-} mice*) and identified a significant increase in total brain area, with significantly enlarged cerebral cortices and corpus callosum structures (Figure 4.5F,G). This increased brain size was accompanied by an increase in cell number (Appendix 4, Table A4.1). Together these results indicate that $Kptn^{-/-}$ mice exhibit megalencephaly affecting most cortical regions with skull overgrowth and deformation, consistent with macrocephaly and frontal bossing seen in *KPTN*-related disorder patients.
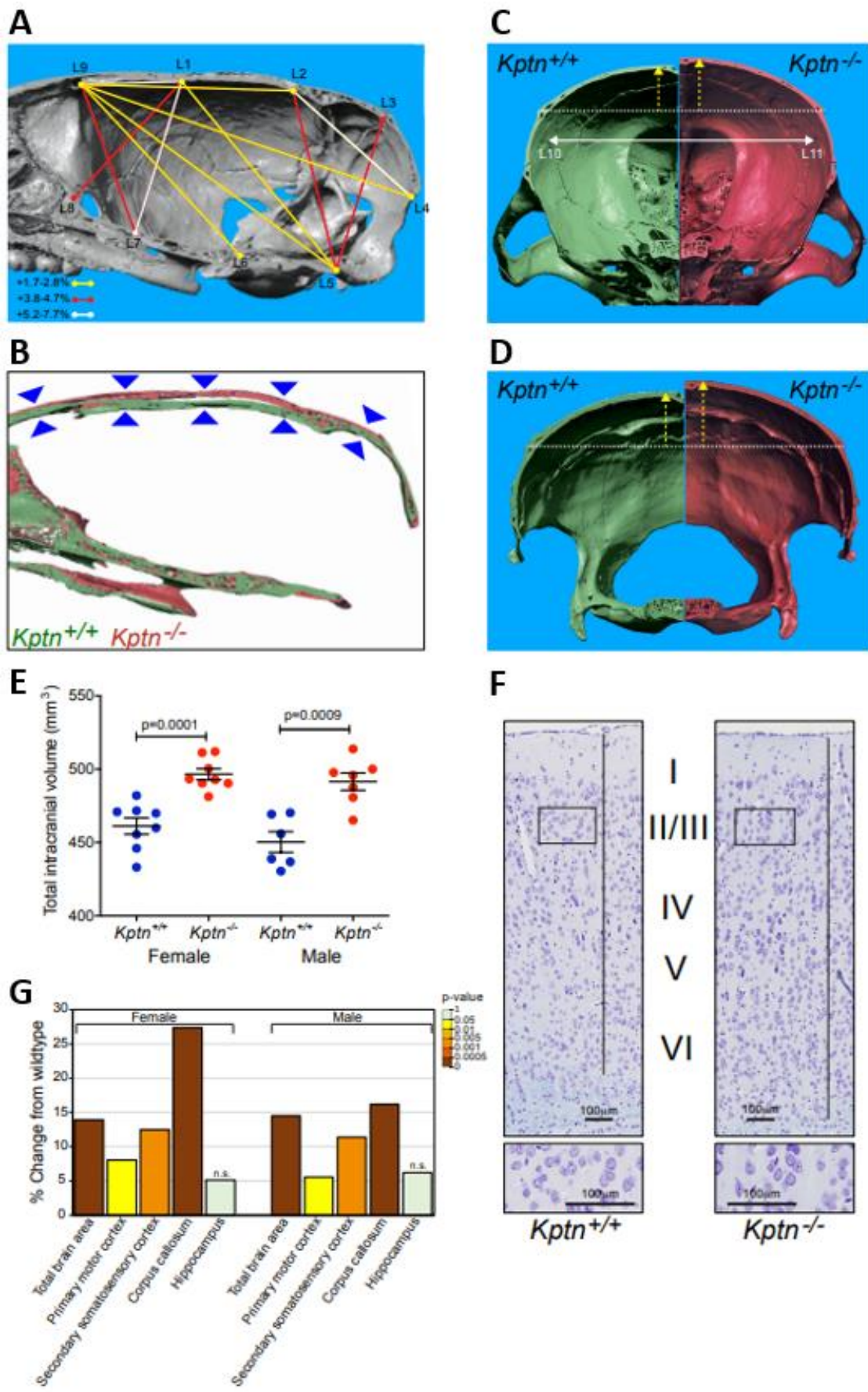
Figure 4.5: *Kptn⁻ᐟ⁻* **mouse model exhibits skull and brain overgrowth**

Micro-computed tomography reconstructions were collected and analysed on male $Kptn^{-/-}$ mice and $Kptn^{+/+}$ controls (n=5 each). (**A**) Significant increases in inter-landmark distances were observed in height, length, and width of the brain cavity of $Kptn^{-/-}$ animals (1.7-7.7% increase in mean linear distances, $p<0.05$, two-tailed Student's t-test). (**B**) Sagittal sections of 3D reconstructions illustrate the increased skull height along the full rostro-caudal extent of the brain cavity (arrowheads). (**C,D**) Rostral and caudal-facing reconstructions of coronal sections from representative individual 3D reconstructions highlight changes in the dorsal curvature (yellow arrows) of frontal and parietal bones in $Kptn^{-/-}$ mice (green = $Kptn^{+/+}$, red = $Kptn^{-/-}$), and the width of the brain cavity (white arrow, +2.56%, $p= 0.0356$). (**E**) Volumetric measurements of MRI from both male and female $Kptn^{-/-}$ mutants and $Kptn^{+/+}$ controls detect 7.7-9.2% (F/M) increases in total intracranial volume at 16 weeks of age ($p<0.001$, two-tailed Student's t-tests). Values are plotted as mean ± SEM. (**F**) Representative sections from $Kptn^{+/+}$ and $Kptn^{-/-}$ cortices stained with cresyl violet. (**G**) Morphometric analyses of histological sections (as in **F**) of $Kptn^{-/-}$ mutant and $Kptn^{+/+}$ controls plotted as percentage difference of $Kptn^{-/-}$ from wildtype mean, indicating significant increases in total brain area (shown for section 1), cortical (shown for section 1) and corpus callosum thicknesses (section 2) ($p<0.002$ in most measurements as indicated, two-tailed Student's t-tests). The hippocampus was not significantly (n.s.) changed in size ($p>0.2$).Data courtesy of Dr Sebastian Gerety and Dr Binnaz Yelcin.

To assess the developmental trajectory of brain overgrowth in the *Kptn*$^{-/-}$ mouse model, the histomorphology of neuroanatomical structures was compared at birth (Postnatal day P0), P20, and 16 weeks of age (Figure 4.6). There were no neuroanatomical differences between *Kptn*$^{-/-}$ and wildtype mice at P0 in most parameters including the total brain area, although there was a small reduction in hippocampal volume in the *Kptn*$^{-/-}$ mice (which recovers at later ages). Normal brain growth ceases at P20 in wildtype mice, although *Kptn*$^{-/-}$ mice lack significant overgrowth at this stage, with increases in brain growth developing into adulthood with a significant increase in total brain area and motor cortex by 16 weeks (Figure 4.6). This indicates that megalencephaly observed in the *Kptn*$^{-/-}$ mouse model is postnatal, progressive, and results from an extended period of brain growth beyond the normal period.
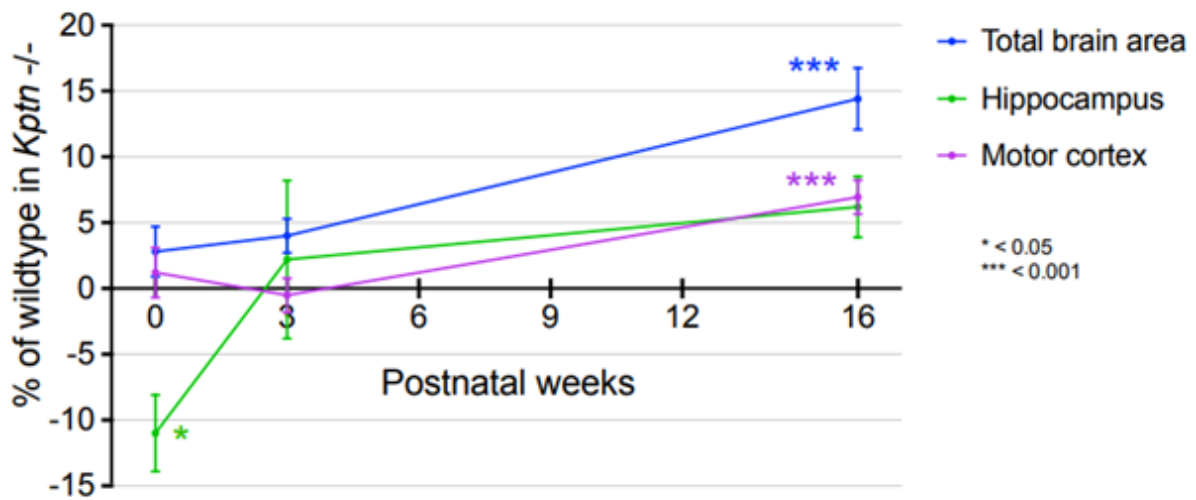
Figure 4.6: *KPTN*^{-/-} mice display postnatal brain overgrowth

Percentage mean change of total brain area, motor cortex and hippocampus in *Kptn*^{-/-} mutant mice at birth, 3 weeks, and 16 weeks postnatally. Significant overgrowth is detected only in 16 week animals ($p<0.001$, two-tailed Student's t-test). Data courtesy of Dr Sebastian Gerety [382].

4.4.2 Clinical features of affected individuals with biallelic *KPTN* variants

We investigated the clinical phenotype of 36 affected individuals (22 male, 14 female) aged 1.5 to 55 years with a confirmed or likely molecular diagnosis of *KPTN*-related disorder. We identified four new patients with Amish heritage and 17 patients with no known links to the Amish. We also revisited the nine Ohio Anabaptist (Amish and Mennonite) patients originally described by our group in 2014 [333] (three are now deceased, brothers A-C, from accidental choking, head injury and pneumonia), and reviewed available clinical data for six patients in four previously published studies [334, 336-338] (see Appendix 4 A4.1.1). Ethnicity, genotype and detailed phenotypic descriptions for newly identified and previously published patients with *KPTN*-related disorder are shown in Tables 4.2 and 4.3. The clinical features of all 36 affected individuals with *KPTN*-related disorder are summarised in Table 4.4.

PATIENT	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8	Patient 9	Patient 10	Patient 11
Genotype	c.776C>A/ c.776C>A	c.776C>A/ c.776C>A	c.776C>A/ c.776C>A	c.776C>A/ c.776C>A	c.776C>A/ c.714_731dup	c.776C>A/ c.714_731dup	c.714_731dup/ c.714_731dup	c.714_731dup/ c.714_731dup	c.714_731dup/ c.714_731dup	c.714_731dup/ c.714_731dup	c.714_731dup/ c.714_731dup
Predicted protein outcome	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/p.(Met241_Gln246dup)	p.(Ser259*)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)
Gender	M	M	M	M	M	M	F	M	F	M	M
Age at evaluation (years)	55	8	7	25	3	5	10	7	45	38	5
Ethnicity	Ohio Mennonite	Ohio Amish	Ohio Amish	Ohio Amish	Michigan, US	Michigan, US	Virginia, US	Denmark	Netherlands	Netherlands	Netherlands
Growth parameters											
Birth Weight kg (SD)	3.74 (+0.38)	3.71 (+0.32)	3.2 (-0.74)	2.60 (-0.23)	3.36 (-0.41)	3.68 (+0.73)	3.03 (-0.82)	2.9 (-0.97)	1.63 (-4.0)	1.7 (-2.7)	3.96 (+0.82)
Birth OFC cm (SD)	N/K	35.6 (+0.28)	N/K	38 (+1.28)	N/K	N/K	35.6 (+0.86)	36 (+1.1)	N/K	30.5 (-1.7)	N/K
Weight kg (SD)	160 (+5.19)	23.7 (-0.7)	25.8 (+0.54)	78.3 (+0.96)	N/K	18.2 (-0.35)	39.3 (+1.01)	20 (-1.2)	87.1 (+2.68)	N/K	19.2 (+0.25)
Height cm (SD)	160 (-2.44)	125 (-0.57)	123.2 (-0.13)	168 (-1.33)	95.5 (-0.79)	105.7 (-1.1)	152.5 (+2.21)	117.5 (-1.0)	168 (+0.74)	N/K	117 (+1.64)
OFC cm (SD)	60.2 (+1.72)	57 (+1.35)	52.7 (+1.47) 2 yrs	61.5 (+2.48)	52.5 (+0.48)	49.2 (-0.44) 21 mths	57 (+2.59)	56 (+1.56)	56.7 (+0.85)	57.5 (+0.14)	54.4 (+1.1)
Maternal OFC cm (SD)	54 (-1.09)	56.5 (+0.71)	N/K	57 (+1.07)	N/K	N/K	54.5 (-0.73)	59 (+2.51)	N/K	N/K	N/K
Paternal OFC cm (SD)	N/A	58.8 (+0.9)	N/K	59 (+1.02)	N/K	N/K	61 (+2.18)	59 (+1.02)	N/K	N/K	N/K
Development											
Walked (years)	5	1.5	2.25	2.5	1	1.3	1.5	2.25	No delay	Delay	2
Expressive/receptive language disorder	Non-verbal	Y	Non-verbal	Y	Y	Y	Non-verbal	Y	Normal	Normal	Y
Intellectual disability	Severe (IQ 25)	Y	Severe	Moderate	Mild (IQ 77-89)	Mild (IQ 86)	Severe	Moderate	Moderate/ Severe	Moderate/ Severe	Mild (IQ 82)
Neurology											
Childhood hypotonia	N	N	Y	Y	Y	N	Y	Y	N	N	N
Seizures	Onset 1y, GTC, absence	N	Onset 6y, GTC	Onset 6m GTC, complex partial	N Staring spells	N	Onset 5y, GTC	Onset 3y, GTC, absence	Onset 27y, GTC, tonic, complex	Onset 19y, GTC, complex partial	N
Neuroimaging	N/P	N/P	N/P	CT/MRI: megalencephaly	Neonatal USS: NAD	MRI: NAD	MRI: mild ventriculomegaly	MRI: megalencephaly	CT: NAD	CT: NAD	N/P
Behavioural characteristics											
Anxiety	N	Y	N	Y	N	N	Y panic attacks	Y	Y	N	Y
Stereotypy	Y	N	Y	Y	N	N	Y	Y	N	N	Y
Repetitive speech	Y	N	Non-verbal	N	N	N	Non-verbal	N	N	N	Y
Social communication disorder	N	Y	Y	Y	Y	Y	Y	Y	N	N	Y
Hyperactivity	N	Y	N	N	N	N	Y	N	N	Y	N
Craniofacial dysmorphism											
Frontal bossing	Y	Y	Y	Y	Y	N/K	N	Y	N	N	Y
Prominent chin	N	N	N	N	N	N/K	N	Y	N	N	Y
Scaphocephaly	N	Y	N	N	N	N (plagiocephaly)	N	N	N	N	N
Hypertelorism	N	N	N	Y	Y	N	N/K	N	N	N	Y	N
Downslanting palpebral fissures	N	N	N/K	Y	N	N/K	N	N	Y	Y	N
Other facial features	N	Unusual pinna	High palate, bulbous nose	High arched palate, flat nasal bridge, hooded eyelids	Thickened ear antihelices	N/K	N	N	Thick vermillion of lower lip	Thick vermillion of lower lip, overhanging nasal tip	Triangular face
Other	Balance problems, swallowing difficulties	Balance problems, right esotropia	Nil	Recurrent otitis media, polyphagia, generalised weakness	Exotropia, oral apraxia, IgA deficiency, joint hypermobility	Amblyopia, otitis media, ventriculoseptal defect (small), oromotor apraxia	Nil	5th finger clinodactyly	short metacarpals 4+5, puffy fingers	Strabismus, acquired brain damage secondary to seizures	High pain threshold, hydrocele, feeding problems, NG feeding, pes planus, otitis media

174

PATIENT	Patient 12	Patient 13	Patient 14	Patient 15	Patient 16	Patient 17	Patient 18	Patient 19	Patient 20	Patient 21
Genotype	c.714_731dup/ c.714_731dup	c.714_731dup/ c.714_731dup	c.980T>G/ c.714_731dup	c.58_66del/ c.714_731dup	c.58_66del/ c.714_731dup	c.395-2A>C/ c.714_731dup	c.394+1G>A/ c.714_731dup	c.597_598dup/ c.597_598dup	c.863G>A/ c.1129G>A	c.863G>A/ c.1129G>A
Predicted protein outcome	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Met241_Gln246dup)/p.(Met241_Gln246dup)	p.(Leu327Arg)/p.(Met241_Gln246dup)	p.(Thr20_Phe22del)/p.(Met241_Gln246dup)	p.(Thr20_Phe22del)/p.(Met241_Gln246dup)	p.?/p.(Met241_Gln246dup)	p.?/p.(Met241_Gln246dup)	p.(Ser200Ilefs*55)/p.(Ser200Ilefs*55)	p.(Arg288Gln)/p.(Asp377Asn)	p.(Arg288Gln)/p.(Asp377Asn)
Gender	F	F	M	F	F	M	M	M	F	M
Age at evaluation (years)	4	7	2	8	5	6	1.5	2	11	14
Ethnicity	Newcastle, UK	Newcastle, UK	Netherlands	Leicester, UK	Leicester, UK	London, UK	New York, US	Liverpool, UK	Canada	Canada
Growth parameters										
Birth Weight kg (SD)	3.54 (+1.42)	3.23 (-0.38)	2.34 (+0.3)	3.04 (-0.81)	3.28 (+0.18)	3.03 (-1.09)	N/K	N/K	2.69 (-1.62)	3.4 (-0.32)
Birth OFC cm (SD)	N/K	N/K	31 (-0.3)	N/K	N/K	34 (-0.95)	N/K	N/K	36 (-1.36) 6 weeks	39 (+0.19) 2 mths
Weight kg (SD)	15.6 (-0.31)	22 (-0.29)	12.5 (-0.14)	22.4 (-1.06)	19.8 (-0.04)	19.7 (-0.49)	N/K	9.55 (-2.68)	44.6 (+1.06)	51.0 (+0.2)
Height cm (SD)	98.1 (-0.85)	123.3 (+0.39)	91.5 (+1.26)	120 (-1.5)	110 (-0.8)	115 (-0.3)	N/K	81 (-2.09)	155.3 (+1.62)	170.6 (+0.99)
OFC cm (SD)	53.5 (+2.0)	53.5 (+0.74)	53.2 (+2.15)	56 (+2.39)	52.8 (+0.62)	55 (+1.17)	N/K	51.5 (+0.87)	57.5 (+2.69)	58.2 (+1.41)
Maternal OFC cm (SD)	N/K	N/K	55 (-0.37)	57.5 (+1.43)	57.5 (+1.43)	55 (-0.37)	N/K	N/K	58 (+1.79)	58 (+1.79)
Paternal OFC cm (SD)	N/K	N/K	61 (+2.18)	55.5 (-1.03)	55.5 (-1.03)	58 (+0.43)	N/K	N/K	57.2 (-0.04)	57.2 (-0.04)
Development										
Walked (years)	2	2.75	2	1.5	2	1.5	Delay	Delay	1.3	1.25
Expressive/receptive language disorder	Y	Y	Y	Y	Y	Y	N/K	Y	Normal	Normal
Intellectual disability	Moderate	Moderate/ Severe	Mild/ moderate	Moderate	Mild/ moderate	Y	Y	Y	Mild	Mild
Neurology										
Childhood hypotonia	Y	Y	Y	N	N	Y	Y	Y	N	N
Seizures	N	N	N	N	N	N	N	Onset 3y, GTC	N	N
Neuroimaging	N/P	MRI: NAD	N/P	N/P	N/P	MRI: ventriculomegaly	MRI: NAD	MRI: NAD	MRI: pineal cyst, megalencephaly	MRI: pineal cyst, megalencephaly
Behavioural characteristics										
Anxiety	N	N	Y	N/K	N/K	Y	N/K	N/K	Y	N (phobias)
Stereotypy	N	N	Y	N/K	N/K	Y	N/K	N/K	N	Y (tics)
Repetitive speech	N	N	N	N/K	N/K	N	N/K	N/K	N	Y
Social communication disorder	N	N	N	Y	N/K	Y	N/K	N/K	Y	Y
Hyperactivity	N	N	N	N/K	N/K	Y	N/K	N/K	Y	Y
Craniofacial dysmorphism										
Frontal bossing	Y	Y	Y	Y	N/K	Y	Y	N	N	Y
Prominent chin	N/K	N/K	N/K	N/K	N/K	N	Y	N	N	Y
Scaphocephaly	N/K	N/K	N (dolicocephaly)	N/K	N/K	N	N	N	N	N
Hypertelorism	N/K	N/K	N/K	N/K	N/K	N	Y	N	N	N
Downslanting palpebral fissures	N/K	N/K	Y	N/K	N/K	Y	Y	N	N	N
Other facial features	Depressed nasal root	Depressed nasal root	Long eyelashes, broad, deep nasal bridge	Low-set ears	Thin vermilion border	Flat nasal bridge	Flat nasal bridge	N	N	N
Other	Recurrent ketotic hypoglycaemia with intercurrent illness, delayed closure of anterior fontanelle	Recurrent ketotic hypoglycaemia with intercurrent illness	Hyperlaxity of fingers	Otitis media	Recurrent pneumonia, Neutropenia, Coeliac disease, Hearing impairment	Hypermobility, Large anterior fontanelle, Umbilical hernia, Hypothyroidism	White-Sutton Syndrome	Idiopathic ketotic hypoglycaemia, pectus carinatum, horizontal nystagmus, poor feeding	High astigmatism, strabismus, otitis media, dyslexia, pre-eclampsia in pregnancy	Otitis media, pectus excavatum, dyslexia, dysgraphia

Table 4.2: Clinical features of newly identified individuals with biallelic *KPTN* variants

Height, weight and OFC Z-scores were calculated using a Microsoft Excel add-in to access growth references based on the LMS method using a reference European population [54, 245, 393]. Abbreviations: SD; standard deviation, Y; feature present, N; feature absent, N/K; not known, N/P; not performed, M; male, F; female, GTCS; generalised tonic clonic seizure, y; years, m; months, OFC; occipitofrontal circumference, IQ; intelligence quotient, MRI; magnetic resonance imaging, CT; computed tomography, USS; ultrasound, IgA; immunoglobulin A, NAD; no abnormalities detected.

PATIENT	Patient A	Patient B	Patient C	Patient D	Patient E	Patient F	Patient G	Patient H	Patient I
Published	Baple et al IX:3	Baple et al IX:4	Baple et al IX:5	Baple et al IX:6	Baple et al IX:8	Baple et al IX:9	Baple et al IX:10	Baple et al X:1	Baple et al X:2
Genotype	c.776C>A/c.776C>A	c.776C>A/c.776C>A	c.776C>A/c.776C>A	c.776C>A/c.776C>A	c.776C>A/ c.714_731dup	c.776C>A/ c.714_731dup	c.776C>A/ c.714_731dup	c.776C>A/ c.714_731dup	c.776C>A/ c.714_731dup
Predicted protein outcome	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/ p.(Ser259*)	p.(Ser259*)/p.(Met2 41_Gln246dup)	p.(Ser259*)/p.(Met2 41_Gln246dup)	p.(Ser259*)/p.(Met2 41_Gln246dup)	p.(Ser259*)/p.(Met2 41_Gln246dup)	p.(Ser259*)/p.(Met2 41_Gln246dup)
Gender	M	M	M	M	M	F	F	M	F
Age at evaluation (years)	28	20	19	16	13	22	24	7	11
Ethnicity	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish	Ohio Amish
Growth parameters									
Birth Weight kg (SD)	2.95 (+0.7)	2.92 (0)	3.46 (+1.3)	1.59 (-0.2)	3.35 (-0.4)	2.89 (-1.2)	3.16 (-0.5)	3.2 (+2.0)	2.75 (+1.1)
Birth OFC cm (SD)	N/K	40.6 (2.9) 6 weeks	35.6 (1.75)	30.5 (-3.21)	51 (3.9) 10 mths	N/K	N/K	36 (+0.49) 2 days	33 (-1.25)
Weight kg (SD)	121.2 (+3.4)	N/K	66.5 (-0.2)	63.1 (+0.1)	51.1 (+ 0.1)	107.9 (+3.6)	82 (+2.2)	23.3 (-0.56)	23.1 (-0.5)
Height cm (SD)	166.7 (-1.6)	N/K	165.1 (-1.8)	169 (-0.7)	161.6 (+0.6)	156.2 (-1.3)	160 (-0.6)	118.5 (-1.42)	123.5 (-0.4)
OFC cm (SD)	62 (+3.0)	N/K	63.5 (+3.6)	62.5 (+3.4)	61 (+3.3)	63 (+5.4)	60 (+3.2)	52.5 (+3.0)	55.4 (+2.1)
Maternal OFC cm (SD)	55.5 (0.1)	55.5 (0.1)	55.5 (0.1)	58 (1.9)	58.5 (2.3)	58.5 (2.3)	58.5 (2.3)	57 (1.2)	57 (1.2)
Paternal OFC cm (SD)	60 (1.7)	60 (1.7)	60 (1.7)	59 (1.2)	N/A	N/A	N/A	59.5 (1.5)	59.5 (1.5)
Development									
Walked (years)	1	1	1.33	1.9	4	3.8	2.4	>2.2	2.2
Expressive/receptive language disorder	Y	Y	Y	Y	Y	Y	Y	Y	Y
Intellectual disability	Moderate	Mild/Moderate	Severe	Moderate	Moderate	Moderate (IQ 45)	Moderate (IQ 42)	Moderate	Moderate (IQ 46)
Neurology									
Childhood hypotonia	N	N	N	Y	Y	Y	Y	Y	Y
Seizures	Onset 3m, GTC, absence	Onset 7y, GTC, absence	Onset 7y, GTC, absence	Onset 16y, GTC	N	N	N	Onset 5y, GTC	Onset 11y, GTC
Neuroimaging	MRI: megalencephaly	N/A	CT: mild ventriculomegaly	N/P	N/A	N/P	N/P	CT: wide metopic suture	CT: megalencephaly
Behavioural characteristics									
Anxiety	Y	N	N	Y	Y	Y	Y	Y	Y Phobias	Y
Stereotypy	N	N	Y	Y	Y	Y	Y	N	N
Repetitive speech	Y	N	N	N	Y	Y	Y	N	N
Social communication disorder	N	N	N	N	N	N	N	N	N
Hyperactivity	N	N	N	Y	N	N	N	N	N
Craniofacial dysmorphism									
Frontal bossing	Y	Y	Y	Y	Y	Y	Y	Y	Y
Prominent chin	Y	Y	Y	Y	N	N	N	N	Y
Scaphocephaly	N	N	N	Y	Y (Sagittal synostosis)	N	Y	No (plagiocephaly)	N
Hypertelorism	N	N	Y	Y	N	N	N	Y	N
Downslanting palpebral fissures	N	Y	Y	N	N	N	N	N	N
Other facial features	N	N	N	N	N	N	N	N	N
Other	5th finger clinodactyly, deceased (accidental)	5th finger clinodactyly, deceased 29 years (head injury)	5th finger clinodactyly, deceased age 30 (pneumonia)	Nil	Nil	Nil	Hepatosplenomegaly, liver cirrhosis, pneumonia	Splenomegaly, fetal finger pads, pneumonia	Fetal finger pads

PATIENT	Patient J	Patient K	Patient L	Patient M	Patient N	Patient O
Published	Pajusalu et al II:1	Pajusalu et al II:2	Thiffault et al II:1	Lucena et al II:1	Miguez et a l II:1	Miguez et al II:2
Genotype	c.665dupA/c.665dupA	c.665dupA/c.665dupA	c.394+1G>A/ c.714_731dup	c.597_598dup/ c.597_598dup	c.597_598dup/ c.597_598dup	c.597_598dup/ c.597_598dup
Predicted protein outcome	p.(Glu222fs*)/ p.(Glu222fs*)	p.(Glu222fs*)/ p.(Glu222fs*)	p.(Met241_Gln246dup)/p.?	p. (Ser200Ilefs*55)/ p.(Ser200Ilefs*55)	p. (Ser200Ilefs*55)/ p.(Ser200Ilefs*55)	p. (Ser200Ilefs*55)/ p.(Ser200Ilefs*55)
Gender	M	F	M	F	F	F
Age at evaluation (years)	32	24	4	5	7	3
Ethnicity	Estonia	Estonia	Kansas, US	Brazil	Spain	Spain
Growth parameters						
Birth Weight kg (SD)	N/K	N/K	3.13 (-0.88)	N/K	N/K	N/K
Birth OFC cm (SD)	34 (-0.95)	37 (+1.99)	N/K	35.5 (+0.78)	34 (-0.44)	N/K
Weight kg (SD)	N/K	N/K	16.8 (0.13)	N/K	N/K	N/K
Height cm (SD)	N/K	N/K	100 (-0.6)	N/K	N/K	N/K
OFC cm (SD)	63 (+4.5)	60 (+4.0)	57 (+4)	55 (+2.78)	55.5 (+2.4)	53.5 (+1.12)
Maternal OFC cm (SD)	N/K	N/K	Normal	N/K	N/K	N/K
Paternal OFC cm (SD)	N/K	N/K	Normal	N/K	N/K	N/K
Development						
Walked (years)	No delay	No delay	2.5	1.33	1.5	1.67
Expressive/receptive language disorder	Y	Y	Non-verbal	Y	Y	Y
Intellectual disability	Moderate	Moderate	Y	Y	Mild/Moderate	Mild/ Moderate
Neurology						
Childhood hypotonia	N/K	N/K	Y	Y	N	N
Seizures	Onset 10m, GTC	N	Onset 3m, GTC	N	N	N
Neuroimaging	CT: megalencephaly	N/K	MRI: ?AVM, megalencephaly	MRI: pineal cyst, cavity of the septum pellucidum	MRI: megalencephaly	MRI: cavum of the septum pellucidum
Behavioural characteristics						
Anxiety	Y	Y	Y	Y	N	N
Stereotypy	Y	N	Y	Y	N	N
Repetitive speech	Y	N	Non-verbal	Y	N	N
Social communication disorder	Y	N	Y	Y	N	N
Hyperactivity	N	N	Y	Y	N	N
Craniofacial dysmorphism						
Frontal bossing	Y	Y	Y	Y	Y	Y
Prominent chin	N	N	Y	N	Y	Y
Scaphocephaly	N	N	N	Y	N	N
Hypertelorism	N	N	N	N	N	N
Downslanting palpebral fissures	Y	N	Y	N	N	N
Other facial features	High palate	High palate	Sunken eyes, thin upper lip	N	Midface hypoplasia	N
Other	Microretrognathia, self-aggression	Microretrognathia	Hepatosplenomegaly, hypoglycaemia, recurrent infections, deceased age 9 (status epilepticus)	Precocious puberty, Hyperprolactinaemia, Hypertrichosis, Dyspraxia	Delayed anterior fontanelle closure	Delayed anterior fontanelle closure, sparse hair

Table 4.3: Clinical features of previously published individuals with biallelic *KPTN* variants

Height, weight and OFC Z-scores were calculated using a Microsoft Excel add-in to access growth references based on the LMS method using a reference European population [54, 245, 393]. Abbreviations: SD; standard deviation, Y; feature present, N; feature absent, N/K; not known, N/P; not performed, M; male, F; female, GTCS; generalised tonic clonic seizure, y; years, m; months, OFC; occipitofrontal circumference, IQ; intelligence quotient, MRI; magnetic resonance imaging, CT; computed tomography, AVM; arteriovenous malformation.

Clinical Feature	Total cases (36)
Sex	F = 14, M = 22
Mean age in years (range)	14.3 (1.5 – 55)
Developmental phenotype	
Intellectual disability	100% (36/36)
- Mild	30% (10/30)
- Moderate	43% (13/30)
- Severe	23% (7/30)
Speech delay	89% (31/35)
- Non-verbal	11% (4/35)
Gross motor delay	50% (18/36)
Neurological features	
Macrocephaly (≥2 SD)	56% (19/34)
Occipitofrontal circumference above the mean	97% (33/34)
Childhood hypotonia	56% (19/34)
Seizures	44% (16/36)
- Generalised tonic-clonic	100% (16/16)
- Absence	31% (5/16)
- Complex partial	19% (3/16)
Brain imaging	61% (22/36)
- Ventriculomegaly	14% (3/22)
Behavioural phenotype	
Anxiety	63% (20/32)
Stereotypy	53% (17/32)
Impaired social interaction	42% (14/33)
Hyperactivity	28% (9/32)
Repetitive speech	28% (9/32)
Craniofacial dysmorphism	
Frontal bossing	85% (29/34)
Prominent chin	40% (12/30)
Downslanting palpebral fissures	37% (11/30)
Hypertelorism	23% (7/30)
Flat nasal bridge	20% (7/35)
Scaphocephaly	16% (5/32)
High arched palate	11% (4/36)
Other features	
Recurrent infections	31% (11/36)
Strabismus/nystagmus	14% (5/36)
Ketotic hypoglycaemia	11% (4/36)
5th finger clinodactyly	11% (4/36)

Table 4.4: Summary of the clinical features of 36 individuals with *KPTN*-related disorder

Clinical features observed in 36 individuals with *KPTN*-related disorder, listed as a percentage of individuals with a characteristic and the number of individuals shown as a fraction in brackets. Abbreviations: M, male; F, female; SD, standard deviation.

Macrocephaly

We set out to clarify whether the macrocephaly observed in patients with *KPTN*-related disorder is also progressive, in order to identify potential postnatal opportunities to intervene in the progression of the disorder. In 20 of the 21 new patients with biallelic pathogenic *KPTN* variants, occipitofrontal circumference (OFC) was above the mean, varying between +0.14 and +2.69 standard deviations (SDS). In total 29% (6/21) of individuals were macrocephalic (defined as ≥2 SD above the mean) with the average OFC +1.37 SD. This is lower than reported in previously published individuals, in whom 93% (14/15) had macrocephaly (average OFC +3.25 SD). A possible explanation for this difference includes an overall younger age at diagnosis of newly identified individuals (with a mean age of 12.8 compared with 15.7 years), another possibility is ascertainment bias. Overall, macrocephaly was identified in 56% (20/36) of all patients with *KPTN*-related disorder. Data on birth OFC was only available for 15/36 individuals, of these none were macrocephalic at birth and 47% (7/15) had an OFC below the mean at birth. Figure 4.7A compares the birth and most recent OFC Z-score measurement for 15 affected individuals, with a highly significant difference (paired t-test; p=0.0001) between these measurements, clearly demonstrating that macrocephaly develops postnatally. Figure 4.7B displays the combined OFC Z-score measurements for 36 individuals with *KPTN*-related disorder and demonstrates a dramatic increase in OFC within the first two years of life; this is further illustrated by Figure 4.7C with measurements from a single male (Patient 14) in the first seven years of life, with OFC crossing two Z-score lines.

In summary, the majority *of KPTN*-related disorder patients are born with OFCs within the normal range, and progress with age to become clinically macrocephalic (>2 SD). These data demonstrate further consistency between the human disorder and our mouse model, which both experience progressive postnatal overgrowth.
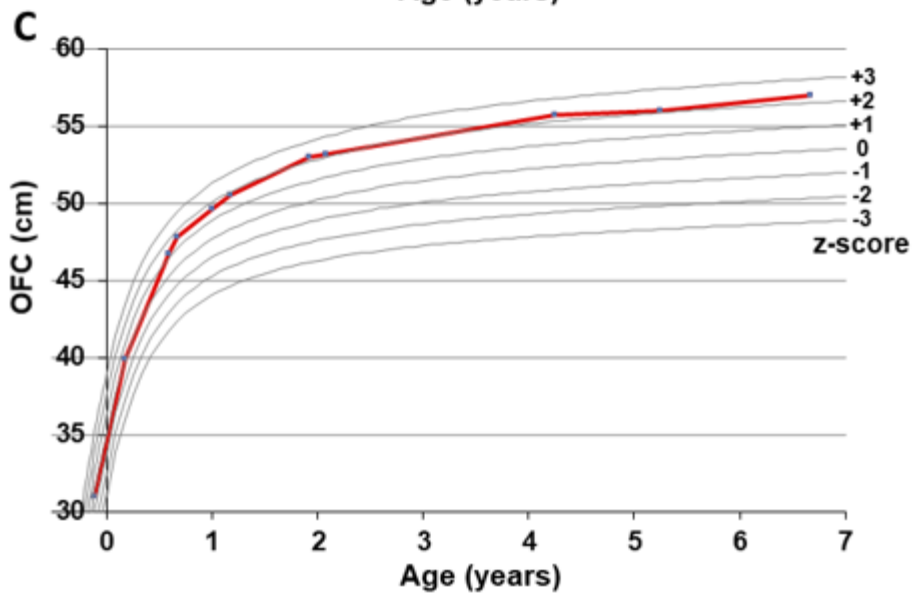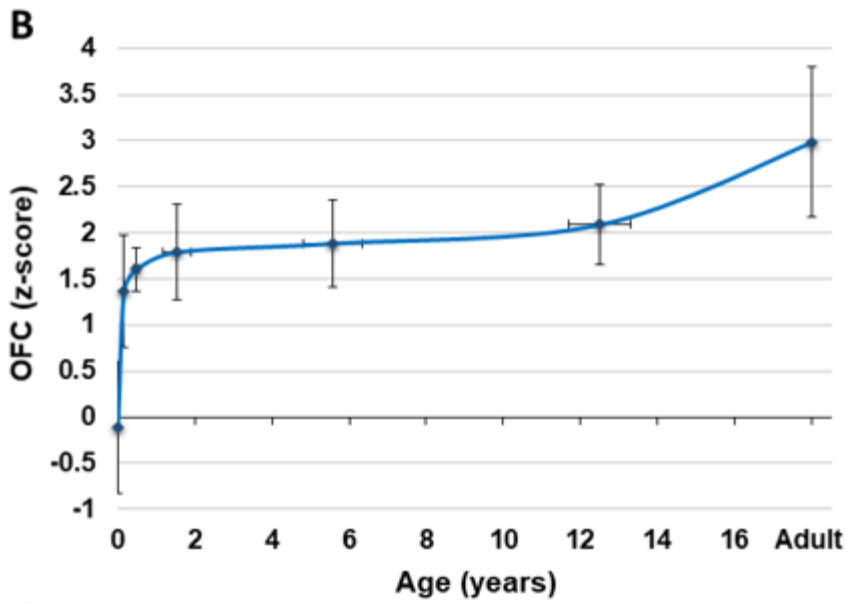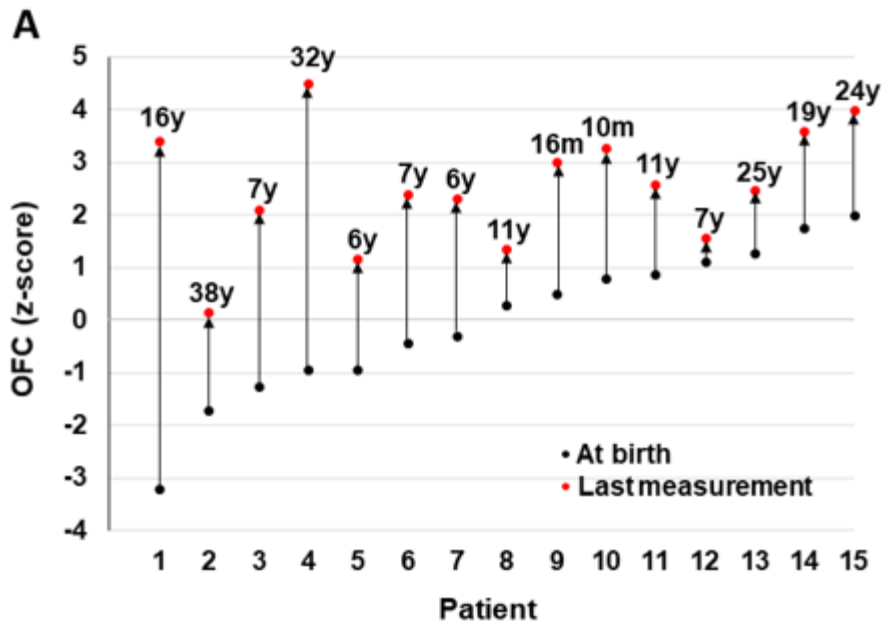
Figure 4.7: Occipitofrontal circumference measurements in individuals with biallelic *KPTN* variants

(**A**) Dot plot demonstrating OFC at birth (black dot) and last measured OFC (red dot) in standard deviations (SD) from the mean for 15 patients with *KPTN*-related disorder. Black arrows show the change in SD for each individual from birth to their last measurement. No individuals had macrocephaly (SD≥2) at birth and all showed an increased in SD from birth to last measurement. Numbers correspond to the following patients in Tables 4.2 and 4.3: 1 = Patient D; 2 = Patient 10; 3 = Patient I, 4 = Patient J; 5 = Patient 17, 6 = Patient N, 7 = Patient 14; 8 = Patient 2; 9 = Patient H; 10 = Patient M; 11 = Patient 7; 12 = Patient 8; 13 = Patient 4; 14 = Patient C; 15 = Patient K. (**B**) Average OFC Z-score measurements for 36 individuals with *KPTN*-related disorder at averaged time points. This demonstrates OFC within the normal range at birth and development of macrocephaly within the first 2 years. Errors bars show standard deviation for the average age and OFC points. (**C**) OFC measurements from birth for Individual 14 demonstrating development of macrocephaly within the first 2 years.

Neuroimaging

Neuroimaging had been performed in 22/36 patients with *KPTN*-related disorder (15 MRI, 6 CT and 1 USS, see Tables 4.2 and 4.3), with 18 patients identified to have generalised megalencephaly with no structural abnormalities. Other findings include mild ventriculomegaly (3/22), pineal cyst (3/22), cavum of the septum pellucidum (2/22), mildly enlarged cortical sulci, wide metopic suture, non-specific white matter hyperintense lesions and supratentorial leukoencephalopathy.

Developmental delay and intellectual disability

Intellectual disability is a universal feature described in all individuals with *KPTN*-related disorder (Tables 4.2 and 4.3) and varies from mild to severe (Table 4.4), although no genotype-phenotype correlation could be identified; severity of intellectual disability appears to increase with the age and the presence of seizures. Patient H and I (siblings X:1 and X:2) from our previously published cases, were reassessed and found to have moderate intellectual disability (previously reported as mild). Developmental delay is a variable feature (see Table 4.4).

Six Amish individuals with *KPTN*-related disorder (Patients C-G and I, Table 4.3) between the ages of 11 and 29 and six age matched controls from the same population were psychometrically assessed for VCI, PRI, PSI, WMI, FSIQ, list learning and narrative memory (see section 2.2.2.1). We detected significant impairment of cognitive function in all six individuals with *KPTN*-related disorder, with all scores within the impaired range (Z-score <-2) compared with age and population matched controls, except the relative sparing of narrative memory function (Table 4.5, Figure 4.8). Individuals with *KPTN*-related disorder had a mean FSIQ score of 43 (Z-score -3.83, Table 4.5), confirming a moderate level of intellectual disability in these individuals. Performance on the story memory test was within the lower normal range (Z-score -1.07), suggesting these individuals have intact narrative memory function. As narrative memory can be considered to have less of a hippocampal contribution [394], these psychometric test results are consistent with the results found in our mouse model. This cross-species comparison highlights the concordance of our mouse model with the human disorder at a neurocognitive level.

SCORE COMPARISON	Control	Affected	Paired t-test
Gender	M = 0 (0%) F = 6 (100%)	M = 3 (50%) F = 3 (50%)	-
Age at evaluation (years)	15.7	20	P = 0.2624
Population	Ohio Amish	Ohio Amish	-
Neurodevelopment			
Intellectual disability	0/6 (0%)	6/6 (100%)	-
Neuropsychiatric data			
Full scale IQ (Z-score)	84.5 (-1.0)	42.5 (-3.8)	P < 0.0001
Verbal comprehension (Z-score)	80.8 (-1.3)	51.5 (-3.2)	p < 0.0001
Perceptual reasoning (Z-score)	84.7 (-1.0)	49.7 (-3.4)	p < 0.0001
Working memory (Z-score)	107.5 (0.5)	52.2 (-3.2)	p < 0.0001
Processing speed (Z-score)	85.3 (-1.0)	51.8 (-3.2)	p = 0.0002
List learning (Z-score)	9 (-0.3)	2.2 (-2.6)	P = 0.0133
Narrative memory (Z-score)	11.7 (0.6)	6.8 (-1.1)	P = 0.1653

Table 4.5: Results of psychometric tests in a subset of patients with *KPTN*-related disorder

Table of averaged demographic details, phenotype and psychometric score results for six Amish individuals with *KPTN*-related disorder and six age-matched Amish individuals and statistical comparison with p-value of unpaired t-test to compare score results between the two groups, showing value/score and Z-score in brackets. Abbreviations: IQ: intelligence quotient, M: male, F: female.
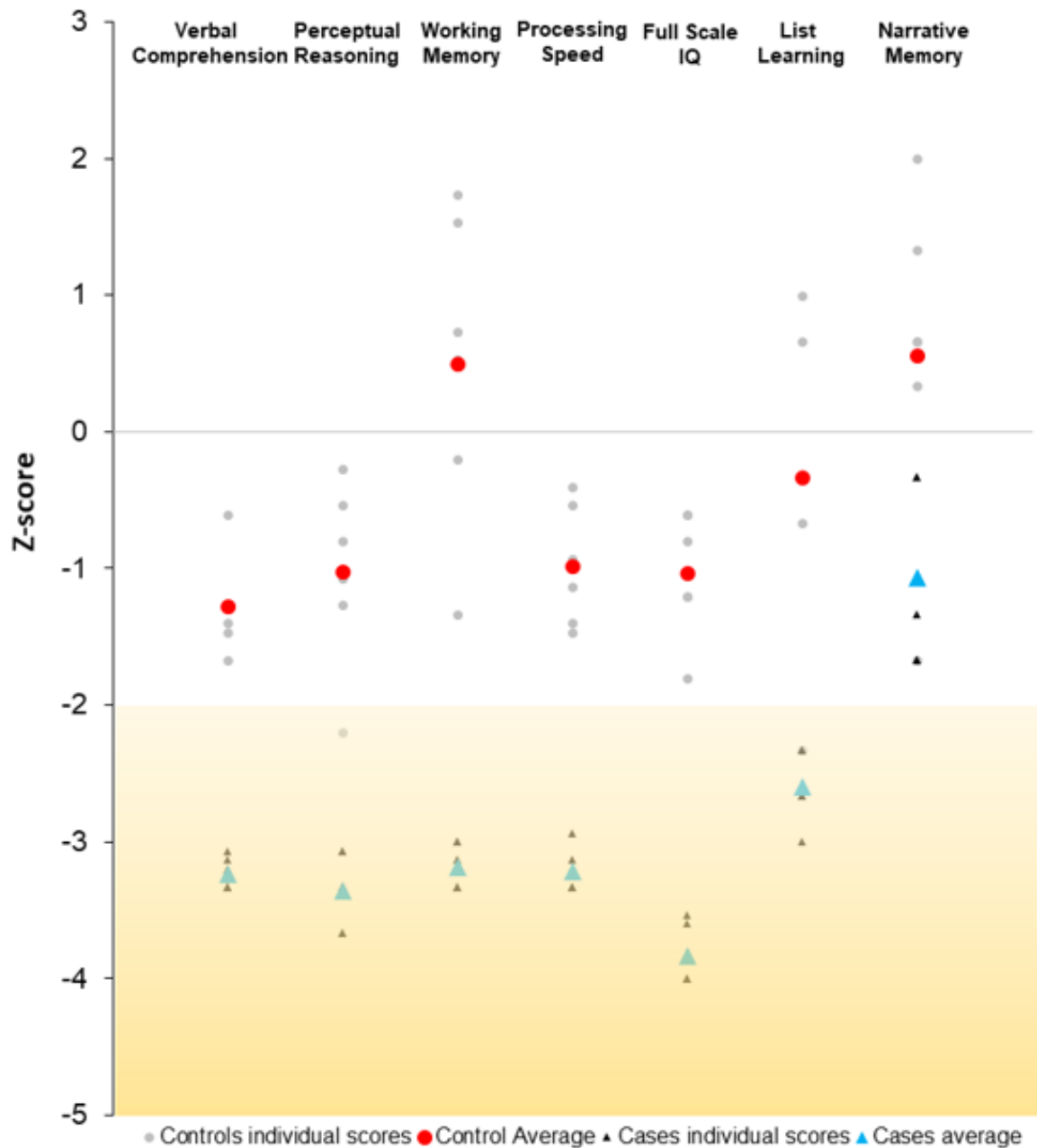
Figure 4.8: Psychometric testing of Amish individuals with *KPTN*-related disorder reveals cognitive sparing

Six Amish individuals between the ages of 11 and 29 with *KTPN*-related disorder (3 female, 3 male), and age-matched population controls, were psychometrically assessed using WISC-IV (6-16 years) or WAIS-IV (over 16 years), to assess cognitive performance in four domains including VCI, PRI, PSI and WMI, combined to generate a FSIQ score. Auditory-verbal recall performance was also assessed by list learning and story memory tests of NEPSY-II (5-16 years) or RBANS-Update (over 12 years). Affected individuals were consistently in the impaired range (Z-score <-2) except in the narrative memory assessment, where probands and controls were not significantly different (p=0.1653).

Neurological features

Seizure phenotype

In the original 2014 description of *KPTN*-related disorder 3/9 patients had a history of seizures [333]. Reassessment of the original Amish pedigree revealed that Patient H and I, aged 6 and 12, had both subsequently developed generalised tonic-clonic seizures (GTCS) which are refractory in Patient H despite polytherapy.

Seizure phenotype observed in newly identified individuals with *KPTN*-related disorder is similar to previously published cases with GTCS in 38% (8/21), with variable age of onset between 6 months and 27 years (Tables 4.2, 4.3 and 4.4); and additional absence or complex partial seizures in some patients (summarised in Table 4.4). Electroencephalogram (EEG) studies were available for two affected individuals (Patients 8 and 9) with similar abnormalities of generalised slowing indicative of diffuse cerebral dysfunction, with multifocal epileptiform discharges and frequent bifrontotemporal spike and slow wave.

Behavioural and psychiatric issues

Behavioural features are a common (85% (28/33)) but variable feature in *KPTN*-related disorder (Tables 4.2 and 4.3) and are summarised in Table 4.4One newly identified individual had tic disorder, not previously described in published cases, although commonly associated with other behavioural features. There have also been anecdotal reports of an increased pain threshold in a number of affected individuals.

Craniofacial dysmorphism

Craniofacial features observed in newly identified patients were consistent with those previously described and include: frontal bossing, scaphocephaly, a prominent chin, small downslanting palpebral fissures, hypertelorism, a broad nasal tip, deep/flat nasal bridge, thick vermillion of the lower lip, high arched palate and mild retrognathia in some individuals. These are illustrated in Figure 4.9A-P and summarised in Tables 4.2, 4.3 and 4.4.

Other clinical features

Other variable clinical features of *KPTN*-related disorder include delayed anterior fontanelle closure (3/36), with raised intracranial pressure requiring shunt insertion in

a single individual aged 2.5 years (Patient 17), 5th finger clinodactyly and fetal finger pads. Recurrent infections were described in 11/36 (31%) patients including pneumonia and otitis media; one individual had neutropenia (Patient 16) and another had IgA deficiency (Patient 5). Ophthalmic features not previously described include strabismus, amblyopia, high astigmatism and horizontal nystagmus. Hepatosplenomegaly was identified in two previously published patients, one with additional liver cirrhosis, and one further individual had splenomegaly only. Endocrine abnormalities include hypoglycaemia, hypothyroidism and hyperprolactinaemia. We describe a further three patients (12, 13 and 19) with idiopathic ketotic hypoglycaemia, affecting 11% of *KPTN*-related disorder patients overall.
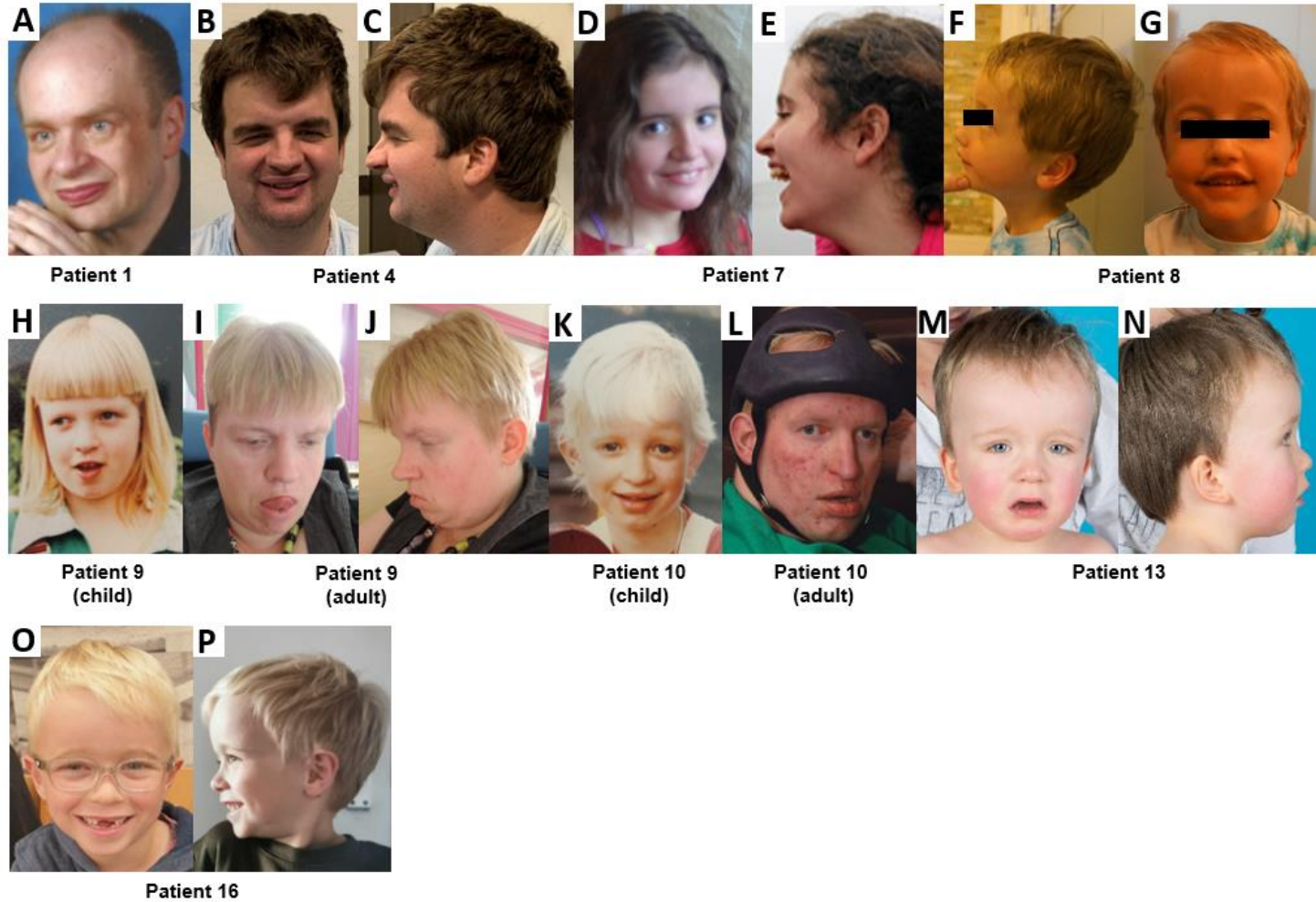
Figure 4.9: *KPTN*-related disorder is a macrocephalic neurodevelopmental disorder with subtle facial dysmorphology

(**A-P**) Facial appearance of nine newly identified patients with *KPTN*-related disorder, identifiable features include: frontal bossing, a prominent chin, small downslanting palpebral fissures, hypertelorism, a broad nasal tip, thick vermillion of the lower lip and mild retrognathia.

4.4.3 Human brain growth is sensitive to KPTN dosage

During clinical data collection from Amish individuals affected by *KPTN*-related disorder, it was observed that a number of their parents presented with larger than average head size. Given that these parents are all heterozygous carriers of pathogenic *KPTN* variants, a systematic analysis of OFC was undertaken in 24 heterozygous parents of identified cases (Tables 4.2 and 4.3). We observed a significant increase in OFC size (mean Z-score 0.856, p=0.000028) in carriers compared to the general population distribution, though infrequently in the macrocephalic range (4 out of 24 had an OFC >2SD, Figure 4.10A). This observation suggests that while *KPTN*-related disorder is an autosomal recessive condition, pathogenic *KPTN* alleles display a heterozygous effect on head size, possibly due to the same mechanisms of mTOR hyperactivation. To test this directly, we differentiated heterozygous *KPTN*$^{+/-}$ loss of function iPSCs into cortical neural precursors. RNAseq on these cells reveals a strong correlation between the gene dysregulation in homozgyous and heterozygous *KPTN* cell lines (pearson r=0.759, p<0.0001, Figure 4.10B). This identified an intermediate mTOR pathway dysregulation in the heterozygous cells, and a widespread disruption to the ribosomal protein gene network, consistent with that observed in the homozygous *KPTN*$^{-/-}$ cell lines (Appendix 4, section A4.2.2, Figure A4.5B,C). These data strongly suggest that human head size is sensitive to *KPTN* gene dosage, likely through its action within the mTOR pathway. Although none of the heterozygous carrier parents had any reported cognitive impairment or history of seizures, further studies will be required to understand whether the brain growth effect is accompanied by any changes in neurological function.
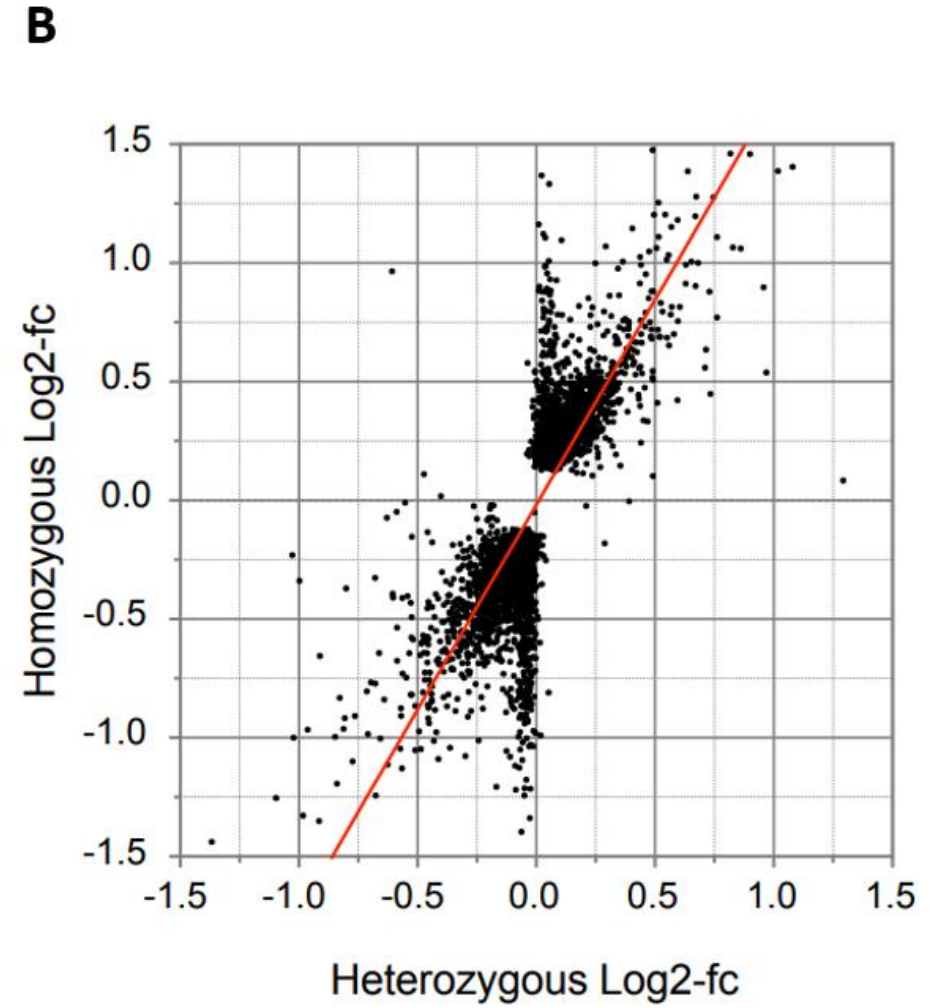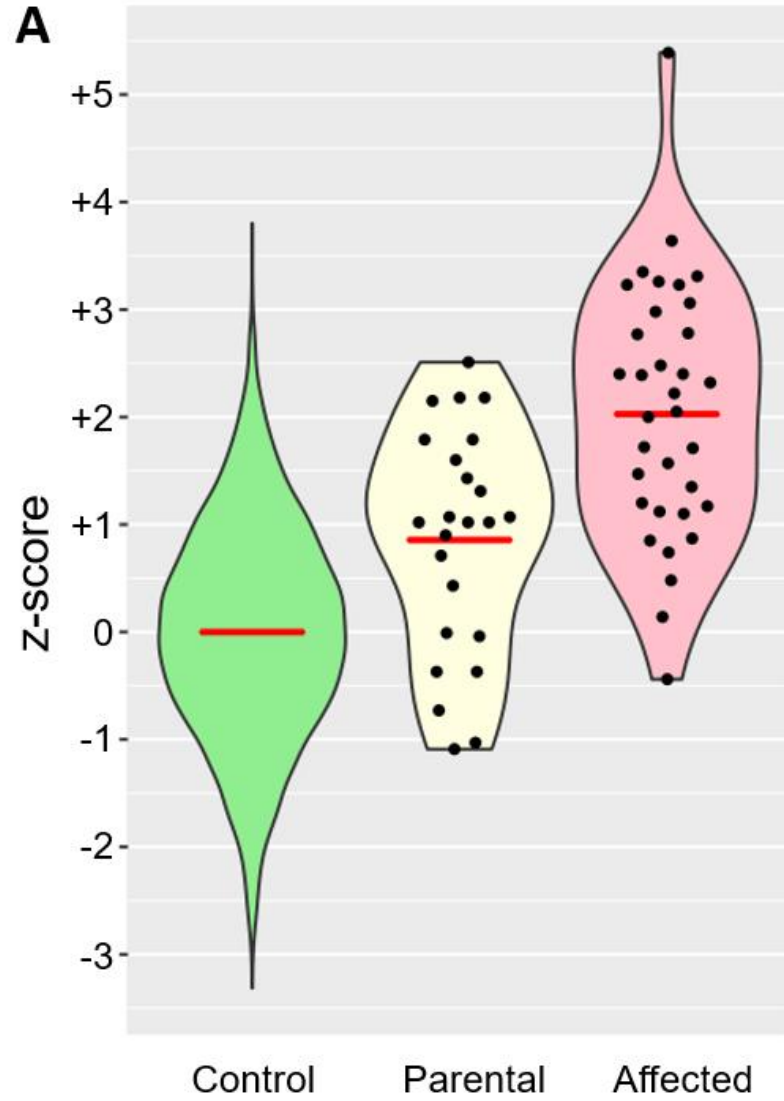
Figure 4.10: OFC measurements in heterozygous parents of *KPTN*-related disorder patients

(**A**) OFC measurements in 24 heterozygous parents (Parental) reveal significant increase in mean head size (mean z-score=0.856, p=0.000028, one sample, 2-sided z-test against wildtype distribution, n=24) compared to normal OFC distribution (Control), and an intermediate effect relative to affected individuals with biallelic *KPTN* variants (Proband, mean z-score=2.027, p<0.00001, one sample, 2-sided z-test against wild type distribution, n=35). (**B**) RNAseq on KPTN LOF NPC models indicates a strong correlation between the gene dysregulation in homozygous and heterozygous KPTN cell lines (pearson r=0.759, p<0.0001).

4.4.4 *Kptn* regulates mTOR signalling *in vivo*

To examine whether KPTN modulates mTORC1 signalling in the *in vivo* model, Western assays were performed on brain tissue from *Kptn*[-/-] and *Kptn*[+/+] mice. Ribosomal protein S6 (RPS6) is a known downstream target of the mTORC1 pathway, whose phosphorylation state is strongly linked to pathway activation [395, 396]. Results in both whole brain and the hippocampus of adult animals reveal a significant increase in mTOR pathway activity in *Kptn*[-/-] mice, as indicated by an increased phosphorylation of RPS6 (p-RPS6, $p<0.05$) (Figure 4.11A). Importantly, mTOR activation is present at P21 ($p=0.0308$, Figure 4.11A), the stage at which the *Kptn*[-/-] mouse model is experiencing brain overgrowth. Immunostaining for p-RPS6 in histological sections of brain tissue highlights the widespread cortical and hippocampal activation of mTOR activity in *Kptn*[-/-] mice (Figure 4.11B). To confirm that the increase in p-RPS6 caused by loss of *Kptn* is mTOR dependent, we treated a cohort of mice with the mTOR inhibitor, rapamycin [397-399]. The excess p-RPS6 signal seen in *Kptn*[-/-] animals is significantly reduced after 3 days of treatment, confirming the presence of hyperactive mTOR signalling in this model ($p=0.00188$, Figure 4.11C). These data are consistent with the proposed role of Kptn as a negative regulator of mTORC1 activity and provides *in vivo* evidence of a role for mTOR signalling in *KPTN*-related disorder.

Transcriptional RNA-seq analysis was performed at the Wellcome Sanger Institute to compare gene expression of mTOR pathway regulators, epilepsy-associated genes, markers of neural stem cells/radial glia and essential neurodevelopmental genes between *Kptn*[-/-] and *Kptn*[+/+] animals and are detailed in Appendix 4, section A4.2.2 (Transcriptomic studies). This work identified extensive dysregulation in *Kptn*[-/-] mice of the expression of both positive and negative regulators of the mTOR pathway (Figure A4.4A,B), numerous genes that control electrophysiological homeostasis in brain tissue and regulate key neurodevelopmental process (Figure A4.5C), with significant increases in expression levels of neural stem cell markers.
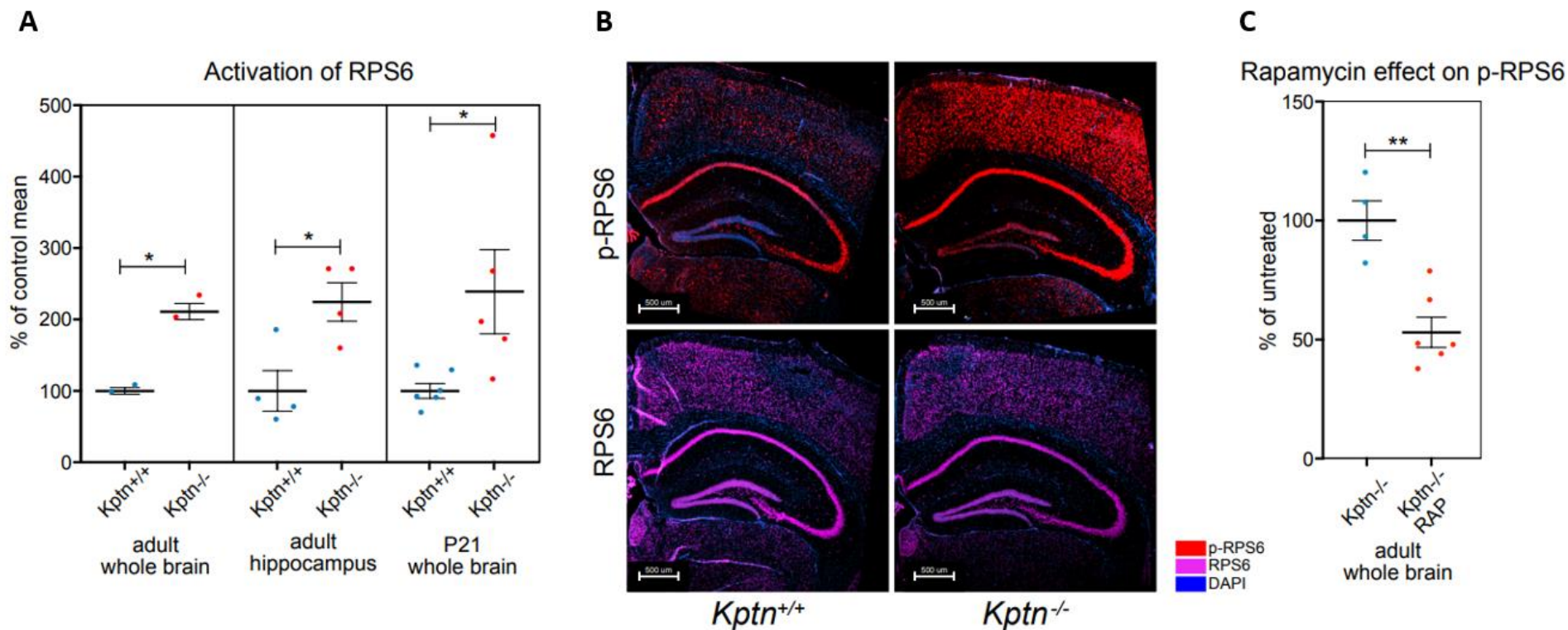
Figure 4.11: Kptn loss-of-function induces changes in mTOR signalling corrected by rapamycin

(**A**) Quantification of the phosphorylation of Ribosomal Protein S6 (RPS6) in *Kptn*[+/+] controls (blue) and *Kptn*[-/-] mutant animals indicate significant increases in mTOR signalling in adult whole brain (p=0.0119, n=2 per genotype), hippocampus (p=0.0189, n=4 per genotype), and in juvenile mice at P21 (p=0.0308, *Kptn*[+/+], n=6; *Kptn*[-/-], n=5), measured as the ratio of phosphorylated RPS6 (p-RPS6) to total RPS6 signal (units as % of control mean). All p-values are from two-tailed Student's t-test). (**B**) Immunostaining for RPS6 and its phosphorylated form (RPS6 and p-RPS6 respectively, red) reveal widespread cortical and hippocampal activation of the mTOR pathway in *Kptn*[-/-] mice (nuclei stained with DAPI, blue). (**C**) Quantification of the phosphorylation of RPS6 in *Kptn*[-/-] animals after 3 days

of treatment with vehicle ($Kptn^{-/-}$, blue) or rapamycin ($Kptn^{-/-}$ RAP, red) indicates a significant reduction in p-RPS6 levels upon treatment (p=0.00188, $Kptn^{+/+}$ n=4; $Kptn^{-/-}$ n=6, two-tailed Student's t-test). Values are plotted as mean ± SEM. Data courtesy of Prof Peter Crino.

4.4.5 *Kptn*-/- cell lines show increased cell size and mTOR activation

Cell line models of *KPTN*-related disorder were generated for further functional studies in mouse neural crest derived Neuro 2A (N2A) cells and human SH-SY5Y neuroblastoma and KOLF2C1 iPSC cell lines. CRISPR/Cas9 gRNAs (plasmid with GFP reporter) were generated and validated to achieve Kptn knockout (gRNAs targeting exons 4/5 or exon 8), aiming to model the known nonsense mutation p.(Ser259*) in humans (see A4.1.3 Molecular studies). We differentiated our iPSC models into PAX6-positive cortical neural precursor cells (NPCs) (Figure A4.5A) [400]. The functional consequence of frameshifting mutations in edited cells was characterised by RNAseq, revealing a dose-dependent reduction in detectable *KPTN* transcript reads in heterozygous (41% reduction) and homozygous (89% reduction) mutant NPC lines, likely reflecting robust nonsense mediated decay of the mutant transcripts. *Kptn* sequencing in edited cells revealed substantive cuts throughout *Kptn* and Western blotting revealed loss of Kptn protein expression (Figure 4.12A). Phosphorylation levels of mTOR target ribosomal protein S6 (PS6) were found to be increased (Figure 4.12A) suggesting an increased activation of mTOR in *Kptn*-/- N2A cells. Cell morphological analysis of *KPTN*-/- SY5Y cells identified an increase in cell soma size (Figure 4.12B), which likely contributes to the observed megalencephaly phenotype.

Transcriptional studies with RNA-seq analysis were performed to compare gene expression of mTOR pathway regulators, epilepsy-associated genes and essential neurodevelopmental genes in *KPTN*-/- human NPCs and are detailed in Appendix 4 section A4.2.2 (Transcriptomic studies) and Figure A4.5. Consistent with the *Kptn*-/- mouse this shows significant dysregulation of gene expression within these groups.
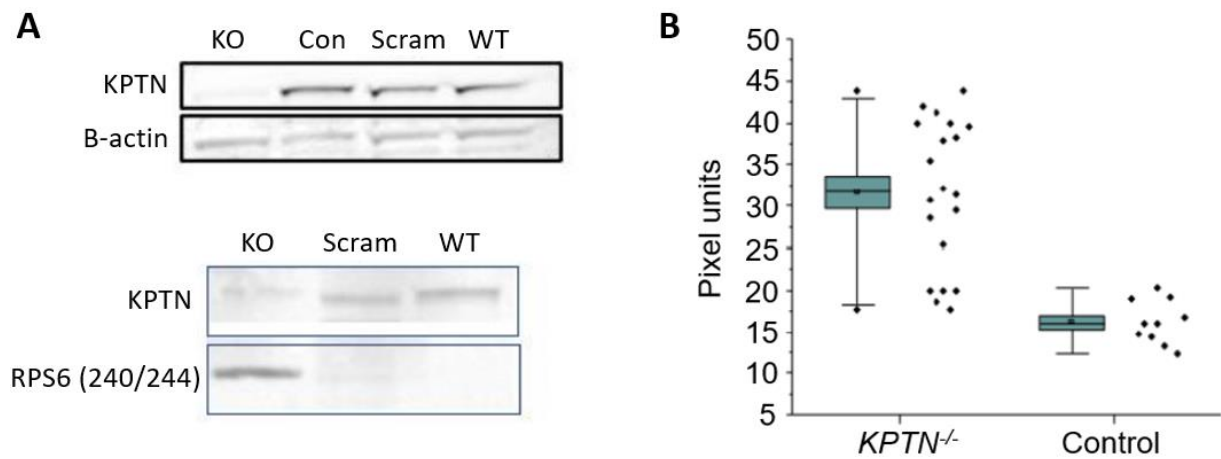
Figure 4.12: CRISPR/Cas9 KO of *Kptn* in N2A cells

(**A**) Top, complete absence of detectable KPTN in *KPTN^-/-* N2A cells. Bottom, near loss of *Kptn* in KO cells, with increased RPS6 levels. (**B**) Increased cell soma size (pixel units) in *Kptn^-/-* SH-SY5Y cells (left) compared with *Kptn^+/+* controls (right). Abbreviations: KO: gRNA *KPTN* KO; Con: control, no effect of gRNA, Scram: scramble on *Kptn,* WT: wildtype. Data courtesy of Professor Peter Crino.

4.4.6 Genetic studies identify ten pathogenic *KPTN* variants

Table 4.6 lists all identified *KPTN* variants associated with *KPTN*-related disorder to date. Of the 21 affected individuals identified in this study, 18 (four Amish and 14 non-Amish individuals) were identified to have segregation of at least one of the two *KPTN* variants (NM_007059.4: Chr19:g.47479874G>T, c.776C>A; p.(Ser259*) and Chr19:g.47479918G>GACCGACCACATCTGCAGA, c.714_731dup; p.(Met241_Gln246dup) [hg38]) originally identified in the Amish community (Figure 4.13A) [333] (Tables 4.2 and 4.3). All variants segregated appropriately for an autosomal recessive disorder within each family (Figure 4.13), were absent or present at low frequency in gnomADv2.1.1/v3.1.2 and assessed to be pathogenic or likely pathogenic according to ACMG variant classification criteria [401] (Table 4.6). Figure 4.13B depicts the distribution of the gene variants within the *KPTN* gene and protein. All three identified missense variants are located within the integrin alpha domain, are predicted to be deleterious by multiple *in silico* pathogenicity prediction tools and are highly conserved (Table 4.6, Figure 4.13B).

Identification of other genetic variants that we were not able to exclude as possibly contributing to the phenotype, include a likely benign small paternally inherited CNV 17q25.1 duplication identified by microarray analysis of Patient 5. WES studies of Patient 18 identified an additional *de novo* likely pathogenic *POGZ* variant (NM_015100.4: c.1523+2T>A; p.?) with a concurrent diagnosis of White-Sutton syndrome.

Within the Amish population the p.(Ser259*) variant was present in a database of >10,000 Amish control individuals (Anabaptist Variant Server), with an allele frequency of 0.0125. While the p.(Met241_Gln246dup) variant was not listed in the Amish Variant Server due to the database configuration parameters, it is present in our in-house Ohio Anabaptist database at an allele frequency of 0.0068 (3/220 controls). These allele frequencies are notably higher than allele frequencies in the Genome Aggregate Database (gnomAD v2.1.1/v3.1.2) (Table 4.6), consistent with allele enrichment (and increased frequency of disease) in some Amish and Mennonite communities [333]. Importantly, there are no homozygous *KPTN* LOF variants listed in gnomAD.

A

Family 2 — Patient 3
Family 3 — Patient 4
Family 4 — Patients 5/6
Family 5 — Patient 7
Family 6 — Patient 8
Family 7 — Patients 9/10
Family 8 — Patient 11
Family 9 — Patients 12/13
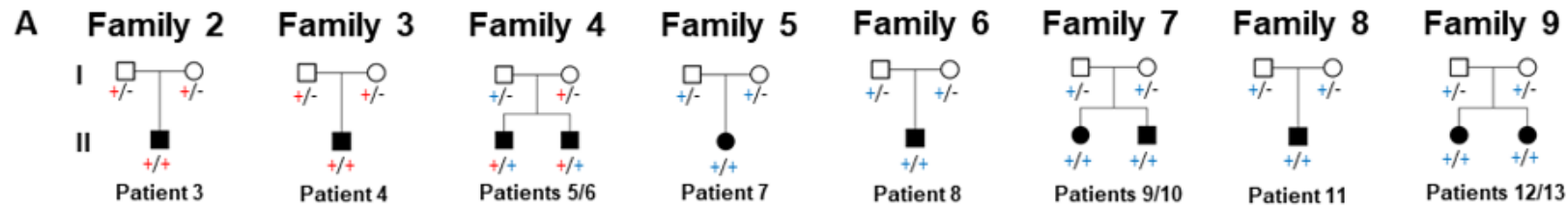
+ c.776C>A; p.(Ser259*)
+ c.714_731dup; p.(Met241_Gln246dup)

Family 10 — Patient 14
Family 11 — Patients 15/16
Family 12 — Patient 17
Family 13 — Patient 18
Family 14 — Patient 19
Family 15 — Patients 20/21
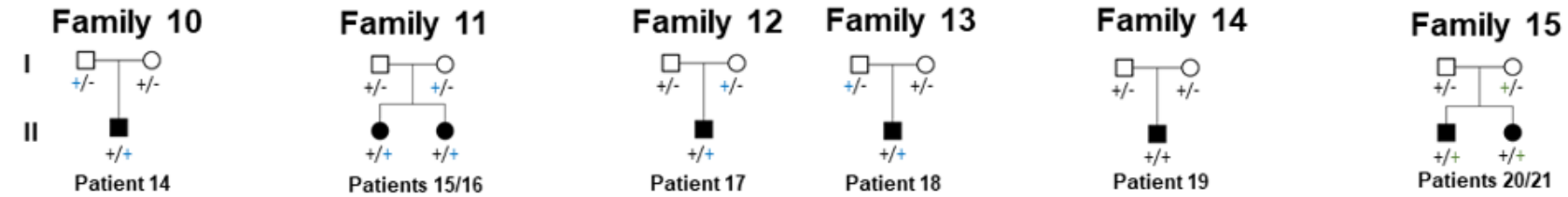
+ c.980T>G; p.(Leu327Arg)
+ c.58_66del; p.(Thr20_Phe22del)
+ c.395-2A>C; p.?
+ c.394+1G>A; p.?
+ c.597_598dup; p.(Ser200Ilefs*55)
+ c.863G>A; p.(Arg288Gln)
+ c.1129G>A; p.(Asp377Asn)

B

Previously described
Present study

p.(Thr20_Phe22del)
c.394+1G>A; p.?
c.395-2A>C; p.?
p.(Ser200Ilefs*55)
p.(Glu222fs*)
p.(Met241_Gln246dup)
p.(Ser259*)
p.(Arg288Gln)
p.(Leu327Arg)
p.(Asp377Asn)

KPTN
NM_007059.4

Integrin alpha domain

0   285   388   417
Q9Y664

Human   VVYRDLL TDVDLDG LLAMAHV
Rhesus macaque VVYRDLL TDVDLDG LLAMAHV
Dog   VVYRDLL TDIDLDG LLAMAHV
Cow   VVYRDLL TDIDLDG LLAMAHV
Mouse   VVYWDLL TDVDLDG LLAMAHV
Zebrafish   VVYRDVE VDLDFDG LLSLIYL

200

Figure 4.13: Pedigrees and *KPTN* variants identified in affected individuals with *KPTN*-related disorder

(**A**) Family pedigrees of newly identified individuals with *KPTN*-related disorder (Family 1 is detailed in Figure 4.1). Genotype segregation is shown under each individual '+', indicates the presence of a *KPTN* variant, '-', wild type allele. Founder variants are identified by red '+' c.776C>A; p.(Ser259*) and blue '+' c.714_731dup; p.(Met241_Gln246dup). Novel variants are indicated in red text. (**B**) *KPTN* gene (NM_007059.4) intron and exon structure with dark blue boxes indicating coding exons, white boxes are non-coding exons and black lines indicate introns (not to scale). The grey bar below represents the KPTN protein (Q9Y664), including an integrin alpha domain, with the location of all previously published (black) and newly identified (red) pathogenic *KPTN* variants. Multi-species alignment is shown for missense variants within the Integrin alpha domain (Families 10 and 15) with conservation of the molecular region encompassing each variant.

CASES	NM_007059.4 (KPTN)	GRCh38 (hg38)	gnomAD v2.1.1 HET	gnomAD v3.1.2 HET	gnomAD HOM	gnomAD All AF	SIFT	Provean	Polyphen-2	REVEL	Splice AI	ACMG Class
Previously published:												
Baple et al	c.714_731 dup p.(Met241_Gln246dup)	Chr19:g.47479918 G>GACCGACCA CATCTGCAGA	135	72	0	0.0004838	*NA*					Pathogenic
Baple et al	c.776C>A p.(Ser259*)	Chr19:g.47479874 G>T	13	15	0	0.0000507	*NA*					Pathogenic
Pajusalu et al	c.665dupA p.(Glu222fs*)	Chr19:g.47480341 C>CT	2	1	0	0.000064	*NA*					Pathogenic
Thiffault et al	c.394+1G>A p.?	Chr19:g.47483294 C>T	19	9	0	0.0000677	*NA*				0.98	Pathogenic
Lucena et al Miguez et al	c.597_598dup p.(Ser200Ilefs*55)	Chr19:g.47480761 C>CTA	29	14	0	0.0001032	*NA*					Pathogenic
New patients:												
Patient 14	c.980T>G p.(Leu327Arg)	Chr19:g.47476822 A>C	1	0	0	0.000005	0.001	-5.63	0.98	0.763	0.00	Likely pathogenic
Patients 15 and 16	c.58_66del p.(Thr20_Phe22del)	Chr19:g.47484094 AGAAGCGCGT>A	0	0	0	0	*NA*					Likely pathogenic
Patient 17	c.395-2A>C p.?	Chr19:g.47483217 T>G	0	0	0	0	*NA*				0.97	Pathogenic
Patients 20 and 21	c.863G>A p.(Arg288Gln)	Chr19:g.47477706 C>T	4	2	0	0.000016	0.23	-1.82	0.56	0.134	0.86	Likely pathogenic
Patients 20 and 21	c.1129G>A p.(Asp377Asn)	Chr19:g.47476585 C>T	1	0	0	0.000004	0.00	-4.17	0.83	0.402	0.00	VUS

Table 4.6: A summary of all *KPTN* variants identified to date in individuals with *KPTN*-related disorder

Includes data from gnomAD (v2.1.1/v3.1.2) of allele count and variant frequencies, pathogenicity prediction tools and variant classification according to ACMG and UK Association for Clinical Genomic Science (ACGS) guidelines [401, 402]. Abbreviations: VUS; variant of uncertain significance, AC; allele count, NFE; non-Finnish European, HET; heterozygous, HOM; homozygous, AF; allele frequency, NA; not applicable, ACMG; American College of Medical Genetics.

4.4.6.1 Haplotype analysis to estimate the founder age of pathogenic *KPTN* variants commonly associated with disease

Genome wide SNP genotyping (Illumina HumanCytoSNP-12v2.1) was performed for 12 individuals with the p.(Ser259*) and p.(Met241_Gln246dup) variants originally identified in the Amish, to determine a shared haplotype length and estimated age of each variant (see Appendix 4, A4.1). In three Amish individuals homozygous for the p.(Ser259*) variant a 2.59Mb genomic region encompassing 155 SNPS (bounded by rs184290 and rs7246244) common to all three affected individuals was identified. In three individuals of mixed ancestry homozygous for the p.(Met241_Gln246dup) variant, a haplotype encompassing 23 SNPs spanning 605Kb, bounded by rs184290 and rs11672818, was defined (Figure 4.14).

An ancestral origin of each haplotype was validated through the analysis of five Amish individuals compound heterozygous for each variant (Patients E – I), in each case the haplotype being consistent. Finally, the haplotype surrounding the novel NM_007059.4: g.47484094AGAAGCGCGT>A, c.58_66del; p.(Thr20_Phe22del) [hg38] variant in Patient 15 was imputed using knowledge of the p.(Met241_Gln246dup) haplotype. Haplotype phasing refined the core ancestral haplotype of the p.(Ser259*) alteration to a 1.2Mb genomic region comprising 78 SNPs (Figure 4.14).

Haplotype analysis indicated that the p.(Ser259*) variant likely arose 18.5 generations ago (95% confidence interval (CI) 11.1 - 25.8), equivalent to approximately 220-520 years (assuming 20 year generations). In contrast the p.(Met241_Gln246dup) variant was estimated to be notably older, arising 34.4 generations previously (95% CI 22.6 - 52.6) or approximately 460-1060 years ago. These estimated dates of origin indicate that each mutation likely arose in Europe in community ancestors prior to the departure of the Amish to the USA, a finding which is consistent with the presence of both alleles in European populations in gnomAD (v2.1.1/v3.1.2).

Figure 4.14: Haplotype analysis of founder _KPTN_ variants

(**A**) Common SNP marker haplotypes flanking the _KPTN_ gene defined in individuals affected with _KPTN_-related disorder, associated with the c.776C>A; p.(Ser259*) and c.714_731dup; p.(Met241_Gln246dup) variants (**B,C**) A common haplotype of 1.2Mb and 605Kb was identified for each variant respectively. Support with analysis of SNP data and calculation of mutational age was provided by Dr James Fasham. Abbreviations: SNP; single nucleotide polymorphism.

4.5 Discussion

This study investigates 36 individuals with biallelic *KPTN* gene alterations to provide a comprehensive assessment of the genetic spectrum and clinical phenotype associated with *KPTN*-related disorder. Multiple behavioural, cognitive, morphological, and molecular investigations demonstrate robust concordance between the *Kptn*[-/-] mouse, and numerous aspects of human *KPTN*-related disorder. This work also enables characterisation of the postnatal progressive nature of brain overgrowth in *KTPN*-related disorder, which provides critical insight into disease progression. Additionally further studies of animal and cellular *Kptn*[-/-] knockout models identify that loss of KPTN is associated with hyperactivation of mTOR and dysregulation of mTOR signalling pathways.

Prior to this work, all identified pathogenic variants were considered likely loss of function, or the inframe p.(Met241_Gln246dup) variant. Fourteen new individuals (Patients 5-18) inherited at least one copy of the p.(Met241_Gln246dup) variant allele originally identified in the Amish, with seven individuals (Patients 7-13) being homozygous for this variant. This is not an unexpected finding, as gnomAD v2.1.1/v3.1.2 estimates the overall carrier frequency of this alteration at approximately 1 in 1,034 (Table 4.6), suggesting a live birth rate for homozygotes of ~1 in 4.3 million. Within the European population (Finnish and non-Finnish) the carrier frequency is higher, approximately 1 in 591 with an estimated live birth rate for homozygotes of ~1 in 1.4 million. Interestingly our haplotype data confirms this variant represents an ancestral founder mutation arising ~34.4 generations previously (Figure 4.14) in Europe prior to the migration of the Amish to the USA in the 17th/18th centuries. The p.(Ser259*) variant was identified in a further six individuals and is estimated to have arisen more recently (~18.5 generations ago) (Figure 4.14). Both variants are present at low frequency in European populations but are enriched in the Amish community, likely being introduced by early founders of the community. Four individuals (three previously published) from three separate regions globally were found to be homozygous for the p.(Ser200Ilefs*55) variant, suggesting this represents another founder variant or recurrent mutation hotspot. This study identifies a further five *KPTN* variants in individuals with *KPTN*-related disorder (Table 4.6, Figure 4.13), all have low frequencies in gnomAD and are classified as 'likely pathogenic' or 'pathogenic' according to ACMG criteria.

The identification of 36 individuals with *KPTN*-related disorder (including 21 previously unpublished patients) from 15 distinct geographical regions, including 13 from the Amish community, has allowed us to perform a comprehensive phenotypic assessment of *KPTN*-related disorder and identify a greater variability and spectrum of clinical features than previously described (Tables 4.2, 4.3 and 4.4). Intellectual disability is universally described, although highly variable (mild to severe), and the degree of developmental delay varies from normal speech development to non-verbal individuals and normal motor development to severe delay, although all children with the condition did eventually achieve independent ambulation. Other features include progressive postnatal megalencephaly (usually developing rapidly within the first 2 years of life), seizures (generalised tonic clonic seizures with additional absence and complex partial seizures), a characteristic behavioural phenotype including anxiety, stereotypies, repetitive speech, hyperactivity and impaired social interaction, hypoglycaemia, hepatomegaly, splenomegaly, strabismus, recurrent infections and dysmorphic facial features. Craniofacial dysmorphism is subtle and it is difficult to recognise this condition on the basis of facial dysmorphology alone (Figure 4.9). Importantly, ketotic hypoglycaemia has been identified in infancy in four affected individuals; if undiagnosed hypoglycaemia could contribute to the development of intellectual disability and can be easily treated. The most consistent neuroimaging finding was mild ventriculomegaly in the absence of any structural brain abnormalities, consistent with generalised megalenecephaly.

Interestingly, the *Kptn*-/- mouse model closely recapitulates the human phenotype, displaying striking similarities with clinical features observed in patients with *KPTN*-related disorder. This includes significant macrocephaly with generalised severe and progressive postnatal megalencephaly but no structural brain abnormalities (Figures 4.5 and 4.6) and behavioural abnormalities including a strong anxiety-like phenotype with significantly increased locomotor activity (hyperactivity) when compared to wildtype mice (Figure 4.4). The results of cognitive tests in *Kptn*-/- mice showed impairment of hippocampal-dependent memory with sparing of hippocampal-independent memory (Figure A4.2), which is mirrored in the psychometric studies of a subset of six Amish individuals with *KPTN*-related disorder who showed impairment of all cognitive domains except narrative memory function, likely due to a compensatory effect as this type of memory utilises a broad distribution of memory

networks with relatively less hippocampal involvement [403]. While not directly tested in our mouse model, we did not observe any clinical seizures during the maintenance and testing of *Kptn*^{-/-} mice, although significant dysregulation of epilepsy-associated gene expression was identified in transcriptomic studies of *Kptn*^{-/-} mice (Appendix 4 A4.2.2).

Macrocephaly was previously considered a universal feature of *KPTN*-related disorder (Table 4.3); notably only ~30% of individuals in our new cohort displayed macrocephaly at the time of assessment, although most had an OFC larger than average within +2SD of the mean (Table 4.3). Explanations for this difference include an overall younger age at diagnosis of newly identified individuals (mean age of 12.8 compared with 15.7 years) and several individuals may not have yet fully developed postnatal macrocephaly, a larger number of published patients had truncating variants and ascertainment bias as these individuals were originally identified due to the presence of macrocephaly. A diagnosis of *KPTN*-related disorder should not be excluded in individuals of normal head size, especially in infants as macrocephaly develops postnatally.

All individuals with *KPTN*-related disorder had a normal head size at birth with a dramatic increase in OFC occurring within the first two years of life and progressing into adulthood (Figure 4.7A-C). The developmental trajectory of brain overgrowth was assessed in the *Kptn*^{-/-} mouse model and identified that significant brain overgrowth results from an extended period of brain growth beyond the normal period (P20) with significant increases in total brain area at 16 weeks (Figure 4.6) and confirm that megalencephaly in *KPTN*-related disorder is postnatal and progressive. Further studies in *Kptn*^{-/-} mouse and human iPSC cell lines identified that megalencephaly likely results from an increase in both cell number and the size of cells (Figures 4.12, A4.3). Transcriptomic studies in *Kptn*^{-/-} mice identified significant persistence of progenitor marker genes (for both radial glia and intermediate neuronal precursors) beyond their normal period of expression (P21) (Appendix 4 A4.2.2: Increases in radial glial and intermediate precursor marker expression) suggesting pathological persistence of proliferating progenitor cells and supernumerary differentiation of neural cell types beyond P21, which could contribute to the progressive postnatal brain overgrowth and cranial malformations seen in *KPTN*-related disorder.

Our findings in *KPTN*-related disorder patients have led to our development of suggested clinical management guidelines (summarised in Table 4.7). These are targeted at preventing life-threatening complications, including early identification and optimisation of seizure control and treatment of ketotic hypoglycaemia, as well as optimising psychomotor development and functional level. It is important for families/carers and clinicians to be aware that seizures may develop at any age: two siblings (Patient H and I, Figure 4.1A, Family 1) identified in our original paper had both subsequently developed GTC seizures and two newly reported siblings (Patients 9 and 10) had onset of seizures in adulthood.

Our data further identifies a heterozygous phenotype in parents and siblings of individuals affected with *KPTN*-related disorder, with these individuals having a significantly larger head circumference (Z-score 0.856, Figure 4.10A) than the general population. Anecdotally we have identified an association with psychiatric manifestations in a number of heterozygous individuals in our study families, which include several individuals with severe depression (including one suicide), anxiety and obsessive-compulsive disorder (OCD). These observations suggests that while *KPTN*-related disorder is an autosomal recessive condition, *KPTN* alleles have a heterozygous effect on head size, with reduced KPTN dosage in heterozygous carriers leading to relative head overgrowth. While the role of somatic and germline mutations affecting mTOR signalling in pathological brain overgrowth is well established [364, 404, 405], it is surprising to find an unrecognised carrier phenotype effect of brain overgrowth. Supporting this finding, constraint metrics such as LOEUF (loss-of-function observed/expected upper bound fraction) [406] indicate that *KPTN* LoF alleles are experiencing purifying selection beyond what would be expected from a purely autosomal recessive disorder gene (observed/expected LoF=0.5, LOEUF = 0.81). Further investigation into the phenotypic consequences of carrier state will be needed to clarify this relatively benign effect on head size, and indeed whether natural variation in other mTOR pathway genes is detectably contributing to differences in brain growth (as suggested by Reijnders *et al* (2017) [407]).

Medical care:	Children should be under the care of a general or community paediatrician to monitor their general health and development. Clinical management strategies should be targeted at early identification and treatment of infections, preventing life-threatening complications and optimising psychomotor development.
Seizures:	Children should be under the care of specialist neurology services for careful optimisation of anticonvulsant medications to maintain seizure control and prevent status epilepticus. Early EEG should be carried out in individuals with suspected seizures. Neuroimaging should be considered at diagnosis.
Gastrointestinal:	Individuals should be examined for evidence of hepatomegaly/splenomegaly.
Speech and language:	Speech and language therapy should be commenced at an early stage to maximise neurocognitive outcome.
Motor development:	Physical and occupational therapies should be commenced at an early stage to maximise neurocognitive outcome.
Vision/Hearing:	Regular ophthalmology and audiology assessments are recommended.
Endocrine:	Infants and children should be monitored for hypoglycaemia and hypothyroidism as these are easily treatable.
Behavioural abnormalities:	Behavioural therapy is beneficial in older affected children who have autism spectrum disorder or hyperactivity.
Education:	An assessment of special educational needs should be carried out so that an individualised educational plan can be put in place at school. Some children have behavioural difficulties requiring additional support and one-to-one instruction.

Table 4.7: *KPTN*-related disorder clinical advice and guidelines

These guidelines are based on the most commonly identified features in individuals with *KPTN*-related disorder. There is a wide range in variability of the clinical presentation and individual patients should have a personalised plan to reflect their own clinical features. Abbreviations: EEG; electroencephalogram.

Since the original description of *KPTN*-associated macrocephalic neurodevelopmental disorder in 2014 [333], new findings regarding the function of this molecule have provided important new insight into the likely pathomolecular basis of the disease. KPTN forms a complex with three other molecules (SZT2, ITFG2 and C12orf66) as part of the KICSTOR complex [367] (Figure 4.3), which interacts with GATOR1 (which in turn is inhibited by GATOR2) to regulate the action of Rag GTPases in response to amino acids and glucose [358] that negatively regulates mTORC1 signalling (Figure 4.3). Through this pathway, mutation of *KPTN* with corresponding disruption of the KICSTOR complex would result in hyperactive mTORC1 signalling, which has been shown to underlie several other neurodevelopmental overgrowth disorders with overlapping features, termed mTORopathies (Table 4.1, Figure 4.3) [153, 408-410].

Biallelic LOF variants within the *SZT2* subunit of the KICSTOR complex result in hyperactivation of mTORC1 and have been shown to cause a severe type of autosomal-recessive infantile encephalopathy with distinct neuro-radiological anomalies [356, 376, 411]. This condition has several overlapping features with *KPTN*-related disorder including macrocephaly, intellectual disability, seizures and dysmorphic facial features, suggesting a common pathogenic mechanism. Mutations within GATOR1 subunits (*DEPDC5*, *NPRL2* and *NPRL3)* have also been associated with hyperactivation of mTORC1 and are a well-established cause of autosomal dominant focal epilepsies [375, 378] (Table 4.1). Wolfson *et al* 2017 identified that the KICSTOR complex interacts with GATOR1 through the SZT2 subunit [367]. This suggests that the interaction between KICSTOR and GATOR1 complexes and their inhibition of mTORC1 is crucial in neurodevelopmental processes, with loss of function of either of these complexes resulting in neurodevelopmental disorders with associated brain abnormalities and epilepsy. Disruption of several other proteins within the mTOR pathway have also been associated with similar neurodevelopmental conditions (Figure 4.3, Table 4.1) including biallelic pathogenic variants in *STRADA* that cause PMSE, through removal of AMPK-mediated inhibition of mTORC1. Three further autosomal dominant conditions (tuberose sclerosis complex (TSC), Cowden macrocephaly-autism syndrome and Smith-Kingsmore syndrome) caused by pathogenic variation in TSC1/2, PTEN and mTOR respectively, all have several overlapping features with the condition including macrocephaly, seizures, behavioural features including autism and hyperactivity, recurrent infections and similar

dysmorphic facial features (Table 4.1) and all impact the mTOR pathway through its hyperactivation (Figure 4.3) [351]. Interestingly, Smith Kingsmore syndrome is caused by gain-of-function variants in mTOR leading to similar overactivation [365].

Our studies of mTOR signalling pathways in *Kptn*[-/-] mouse and human iPSC models support the pathomechanism of *KPTN*-related disorder as resulting from aberrant mTOR signalling. Levels of phosphorylation of ribosomal protein S6, a downstream target of mTORC1 strongly linked to mTOR pathway activation [395, 396] were significantly increased in *Kptn*[-/-] mice (Figure 4.11) and treatment with mTOR inhibitor rapamycin significantly reduced this increase in phosphorylation of S6 and levels of mTOR activation. Transcriptomic studies of both *Kptn*[-/-] mouse and human iPSC models also identified significant dysregulation of gene expression of mTOR pathway regulators (Figures A4.3 and A4.4). These findings are consistent with the proposed role of Kptn as a negative regulator of mTORC1 activity and confirm that *KPTN*-related disorder should be considered an 'mTORopathy'.

mTOR inhibitors have been used as a successful treatment in several mTORopathies including TSC, PMSE and Cowden syndromes that have an underlying pathophysiology of mTOR hyperactivation shared with *KPTN*-related disorder. TSC is a well characterised condition [140] (Table 4.1) with features of benign hamartomas of the skin, brain, kidneys, heart, lungs and retina, as well as neurodevelopmental features of seizures, intellectual disability and behavioural/psychiatric problems. mTOR inhibitors have been successfully used to control resistant seizures and reduce tumor volume in TSC patients [371]. Rapamycin treatment resulted in a 41% decrease in intractable seizure frequency [412] and everolimus has had similar success on seizure control and improved other neuropsychiatric features associated with TSC [413-415]. mTOR inhibitors have been trialled in other mTORopathies, including PMSE and Cowden syndromes with improvement in seizure control and other neurocognitive symptoms observed [369, 370, 372-374]. Alike *KPTN*-related disorder, PMSE is an autosomal recessive mTORopathy with overlapping clinical features and Parker *et al* (2013)[369] treated five PMSE patients with rapamycin starting within the first 8 months of life, with remission of seizure activity in all individuals and improvement in receptive language and social interaction levels reported by parents, although larger studies are required to further confirm these findings. Other promising

targeted second and third generation mTOR kinase inhibitors are being developed and may also have potential therapeutic benefit in mTORopathies [416, 417].

Our pre-clinical therapeutic investigation indicates that mTOR dysregulation in *KPTN*-related disorder is rapamycin sensitive (Figure 4.11C), providing a pathway to possible clinical intervention, particularly for patients with treatment resistant seizures. These findings are clearly exciting given the postnatal onset of several features including megalencephaly and seizures in individuals with *KPTN*-related disorder and potentially provide a period in early infancy where therapeutic intervention may prevent the development of these (and other) clinical features and reduce morbidity. The relationship between megalencephaly and cognitive dysfunction in *KPTN*-related disorder is unclear. The brain overgrowth phenotype could be the primary cause of cognitive deficits observed in *KPTN*-related disorder. Alternatively, neurological phenotypes may result from persistent deregulation of mTOR signalling we observe in adult *Kptn*$^{-/-}$ mice, given the known role of mTOR in cognition and memory formation [353]. If so, the morphological changes in *KPTN*-related disorder could simply reflect the effects of *KPTN* variants on brain growth via overactive mTOR signalling, supported by the strikingly similar neurological features of other mTORopathies [365]. The promising treatment outcomes of related mTORopathy mouse models [418-420] provides encouraging evidence supporting the aim of effective therapeutic intervention in *KPTN*-related disorder. As mTOR inhibitors are generally well tolerated [413], the path to clinical intervention in *KPTN*-related disorder may be tantalisingly short, and follow-up experiments using our mouse and cellular models should provide critical information to develop clinical trials in patients with *KPTN*-related disorder.

4.6 Conclusions and future work

This study entails extensive assessments in humans and mice to provide a comprehensive clinical description, define genotype-phenotype correlations, identify the postnatal onset of the megalencephalic aspect and confirm the mTORopathy pathomechanistic basis of *KPTN*-related neurodevelopmental disorder. The work within this chapter forms the basis of two manuscripts currently in the final preparation stages for publication, one consisting of predominantly clinical data and the other incorporating mouse and cell studies (Appendix 6). Initial investigations presented here identify that mTOR inhibitors are a potential therapeutic treatment that have been used successfully in other mTORopathies and are the focus of our future work on *KPTN*-related disorder. This could provide effective treatment for patients with refractory seizures and there is a possible postnatal therapeutic window to prevent the development of progressive clinical features associated with the condition. In the Amish community development of targeted genetic testing panels for newborn screening are currently being established and include the *KPTN* variants identified in this community to allow an opportunity for early diagnosis and intervention for affected individuals, enabling the best possible outcomes from treatment.

This work highlights the strength of combining clinical investigation with characterisation of disease models, allowing a more thorough description of both human and model disease pathophysiology. Using our mouse findings in cognitive function, we were able to subsequently identify areas of cognitive sparing in the Amish patients. Through a careful examination of the developmental trajectory of brain overgrowth in our mouse model, we were able to confirm the progressive nature of *KPTN*-related disorder overgrowth. Conversely, we were able to take clinical observations from apparently unaffected parents and confirm in the cellular models that the heterozygous carriers of *KPTN* LOF mutations likely experience mTOR dysregulation, contributing to a significant increase in head circumference.

Our research group has been successfully awarded major collaborative funding from the National Institutes of Health (NIH) over the next 7 years to take our studies of *KPTN*-related disorder forward, alongside Professor Peter Crino's team (University of Maryland) and I am pleased to be involved in these plans for future research. This

includes studies to robustly confirm our initial findings reported in this thesis, with three main aims of these research studies:

1. To define the effects of KPTN KO *in vitro* on mTOR pathway activation, cell morphology, cell proliferation, and cell motility

2. Analysis of the *Kptn^{-/-}* mouse brain: histopathology, mTOR signalling and hyperexcitability

3. To define the neuro-phenotypic spectrum of *KPTN*-related disorder, with development of a *KPTN*-related disorder patient database

This research will study the effect of *KPTN^{-/-}* in mice and cellular models, with analysis of histopathology, mTOR signalling, cell proliferation and motility and will include further detailed studies of the effects of mTOR inhibitors on the phenotypes of these models. Frequently mouse models of human disorders do not display seizures, due to a higher seizure threshold in mice. Although seizure activity was not observed in *Kptn^{-/-}* mice in the studies described in this thesis, this was not specifically investigated. Our future studies will examine seizure models in *KPTN^{-/-}* mice, through specific protocols for seizure induction and test whether these mice display a reduced seizure threshold and whether mTOR inhibitors have any effects on seizure generation or frequency. Other aspects include work to refine our understanding of the timing and affected cell types during postnatal development, including whether there is persistence of radial glia and intermediate neuronal precursors that may contribute to progressive brain overgrowth and this will contribute to a more informed therapeutic strategy in treating *KPTN*-related disorder and other mTORopathies.

Further studies of patients with *KPTN*-related disorder will involve establishing a worldwide secure registry and natural history database for affected patients and families and we will work closely with The KPTN Alliance (https://kptnalliance.org/) family support group to provide information and support to patients identified with *KPTN*-related disorder. Further studies will be required to confirm the dosage effect on head circumference in heterozygous parents or relatives of individuals with *KPTN*-related disorder and whether this is accompanied by any neurological or psychiatric features. In addition, exploration of whether natural variation in other mTOR pathway genes is detectably contributing to differences in brain growth should be pursued. Following extensive mouse pre-clinical studies, we hope to eventually develop clinical

trials of mTOR inhibitors (for example rapamycin) in patients with *KPTN*-related disorder. This will have two aspects; to assess mTOR inhibitors as a treatment for seizures in children and adults with the condition, and as a potential preventative treatment in infants aimed at slowing or preventing postnatal development of the disorder.

5 IMPROVING KNOWLEDGE OF THE SPECTRUM AND CAUSES OF DISORDERS OF BRAIN GROWTH AND DEVELOPMENT IN COMMUNITIES

5.1 Introduction

An ultra-rare disorder can be defined as a condition affecting fewer than 1 in 50,000 people in the population [421]. Despite genetic causes accounting for a large proportion of neurodevelopmental disorders overall, individually conditions associated with variants in specific genes are often ultra-rare. Monogenic neurodevelopmental disorders can be extremely difficult to study in a global setting due to their rarity, phenotypic, genetic and allelic heterogeneity (Figure 1.3), making these conditions difficult to identify, diagnose and manage, and therefore many patients may experience a challenging 'diagnostic odyssey'. Current diagnostic yields in patients with intellectual disability or GDD is >15% for chromosomal microarray and >35% for exome or genome sequencing [422]. At least 50% of neurodevelopmental disorders have an unknown aetiology, although a large proportion are thought to arise from as yet unrecognised genetic causes [423]. It is commonplace for ultra-rare neurodevelopmental disorders to have very little or no previously published information about a disease associated with variants in a specific gene, rendering it difficult to establish a definitive diagnosis. Providing adequate evidence to confirm a new disease-gene association can be challenging, particularly in view of several recent studies that have refuted previous disease-gene associations with the rise of genomic variant databases such as gnomAD [424-426]. Conclusively identifying and robustly confirming disease-gene associations is crucial to the interpretation of genetic variants and obtaining a precise molecular diagnosis for patients affected by a monogenic neurodevelopmental disorder, and specific guidelines have been developed for assessing the clinical validity of disease-gene associations [123, 427-429].

The study of inherited neurodevelopmental disorders in genetically isolated populations such as the Amish and Pakistani communities described within this thesis, provide an important opportunity to learn about the genetic basis of these conditions whilst also providing healthcare benefits to these families and communities. This

knowledge and improved understanding of the genetic and molecular basis of inherited neurodevelopmental disorders through identification of novel disease genes and variants within these communities may then be applied to other more diverse communities to facilitate clinical benefits for patients worldwide.

This chapter details genetic and clinical data to enable the consolidation of two disorders of brain growth and development due to variants in centrosomal protein-55 kDa (*CEP55*) and inositol polyphosphate-4-phosphatase type IA (*INPP4A)*. Both genes have two and four previously published papers respectively, each entail descriptions of single families in which gene variants were identified, often with limited clinical details and are previously poorly described conditions. The clinical and genetic studies presented here of families with these conditions identified in Amish and Pakistani communities, leading to collaboration with other groups internationally investigating families elsewhere affected by these conditions, enabled the consolidation of both genes as causes of neurodevelopmental disease, and a more accurate delineation of their clinical features.

5.2 An Amish founder variant consolidates disruption of CEP55 as a cause of hydranencephaly and renal dysplasia

5.2.1 Summary

The centrosomal protein 55 kDa (CEP55 (OMIM #610000)) plays a fundamental role in cell cycle regulation and cytokinesis. However, the precise role of CEP55 in human embryonic growth and development is yet to be fully defined. Here a novel homozygous founder frameshift variant in CEP55 present at low frequency in the Amish community, was identified as the cause of disease in two siblings presenting with a lethal fetal disorder. The features of the condition are reminiscent of a Meckel-like syndrome comprising of Potter sequence, hydranencephaly, and cystic dysplastic kidneys. These findings, when considered alongside two recent studies of single families reporting loss of function candidate variants in CEP55, confirm disruption of CEP55 function as a cause of this clinical spectrum and enable us to delineate the cardinal clinical features of this disorder, providing important new insights into early human development.

5.2.2 Introduction

The centrosomal protein-55 kDa (CEP55, OMIM #610000) is a centrosome- and midbody-associated protein that has been shown to play a central role in cell cycle regulation and is recognised as a key protein involved in the abscission process, the final stage of cytokinesis. [430-432]. CEP55 facilitates abscission through the recruitment of two ESCRT (endosomal sorting complex required for transport)-I subunit associated proteins to the midbody: tumor susceptibility gene 101 (TSG101) and apoptosis-linked gene 2 interacting protein X (ALIX) [433] (see Figure 5.1A for bindings sites). The transcribed CEP55 centrosomal protein has three central coiled-coil domains and is expressed at the perinuclear membrane, cytoplasm, and nucleus [431, 434]. Although, the precise role of CEP55 in human embryonic growth and development is yet to be fully defined, recent studies have identified that CEP55 plays a role in neural stem/progenitor cell division/survival and regulation of ciliogenesis [435, 436].

Prior to this study, candidate homozygous nonsense variants in *CEP55* had been identified in five fetuses from two families in association with a lethal disorder, the

features of which include dysplastic kidneys and complex brain malformations [437, 438] (Table 5.1). Homozygous *cep55l* knockout mutant zebrafish display a significant reduction in the size of brain structures, and a decreased number of renal tubules, consistent with the human phenotype [437]. *Cep55* knockout (*Cep55*$^{-/-}$) mouse models have also been shown to display preweaning lethality associated with a wide range of brain abnormalities and polycystic kidneys [435, 436, 439, 440].

This study describes a novel Amish homozygous founder frameshift *CEP55* variant, present at low frequency in the Amish community, in two affected Amish siblings presenting with a lethal fetal phenotype of hydranencephaly and Potter sequence secondary to cystic renal dysplasia with anhydramnios. These findings, considered alongside the two previous studies reporting loss of function candidate variants in *CEP55*, confirm disruption of CEP55 function as a cause of this clinical spectrum and enable us to delineate the cardinal clinical features of this disorder, providing important new insights into early human development.

5.2.3 Methods

Blood and placental samples were collected with informed consent (University of Arizona protocol 10-0050-01) for DNA extraction as described in sections 2.2.1 and 2.4.1. SNP genotyping was performed in a single affected fetus (VI:3) and their non-affected surviving twin (VI:2) as previously described in section 2.4.6. WES analysis (NextSeq 500: Illumina, University of Exeter) was undertaken on a single affected fetus (IV:3) as described in section 2.4.7. Primer design (Appendix 1), PCR and dideoxy sequencing was performed as described in section 2.4.2, to genotype both affected fetuses, unaffected twin sibling and both parents to confirm appropriate segregation. Clinical phenotyping was carried out according to methods detailed in 2.2.2. The *CEP55* variant (NM_018131.4: c.514dup; p.(Ile172Asnfs*17) and associated phenotype data was submitted to ClinVar (www.ncbi.nlm.nih.gov/clinvar, accession SCV000808984).

5.2.4 Results

Clinical findings

An Ohio Amish couple, distantly related through a 4th generation common ancestor (Figure 5.1B), presented in their second pregnancy with dichorionic diamniotic twins,

one male (twin A) and one female (twin B). Antenatal ultrasound scanning (USS) undertaken at 21+2 weeks gestation revealed twin B to have hydranencephaly, a multicystic dysplastic right kidney (the left kidney was not visualised), shortened bowed long bones and anhydramnios. Twin A was found to have bilateral renal pelvis dilatation, but no further abnormalities were detected. A subsequent USS undertaken at 28 weeks gestation to further examine twin B revealed intrauterine growth restriction (IUGR) and pericardial effusion, the bladder and stomach were not identified. The twins were born following an uncomplicated spontaneous vaginal delivery at 42 weeks gestation. Twin B survived for 90 minutes after birth; the clinical features of Potter Sequence with cutaneous syndactyly were documented (Figure 5.1C). Table 5.1 summarises the clinical features of the two affected Amish siblings.

The couple had previously experienced a stillbirth of a male fetus at 42 weeks gestation. The fetus presented with multiple fetal anomalies on antenatal USS at 19 weeks gestation. Anomalies included: IUGR, anhydramnios, hydranencephaly, described as 'hydrocephalus seen throughout the cranium', and bilateral hydronephrosis. The fetal birth weight was 3lb 14oz and features of Potter sequence, bilateral lower limb bowing, talipes and syndactyly were identified (Figure 5.1C). The features observed in both fetuses are reminiscent of a Meckel-like syndrome comprising of Potter sequence, hydranencephaly and cystic dysplastic kidneys. Karyotyping of placental samples was normal but post-mortem (PM) examination was declined for both fetuses.
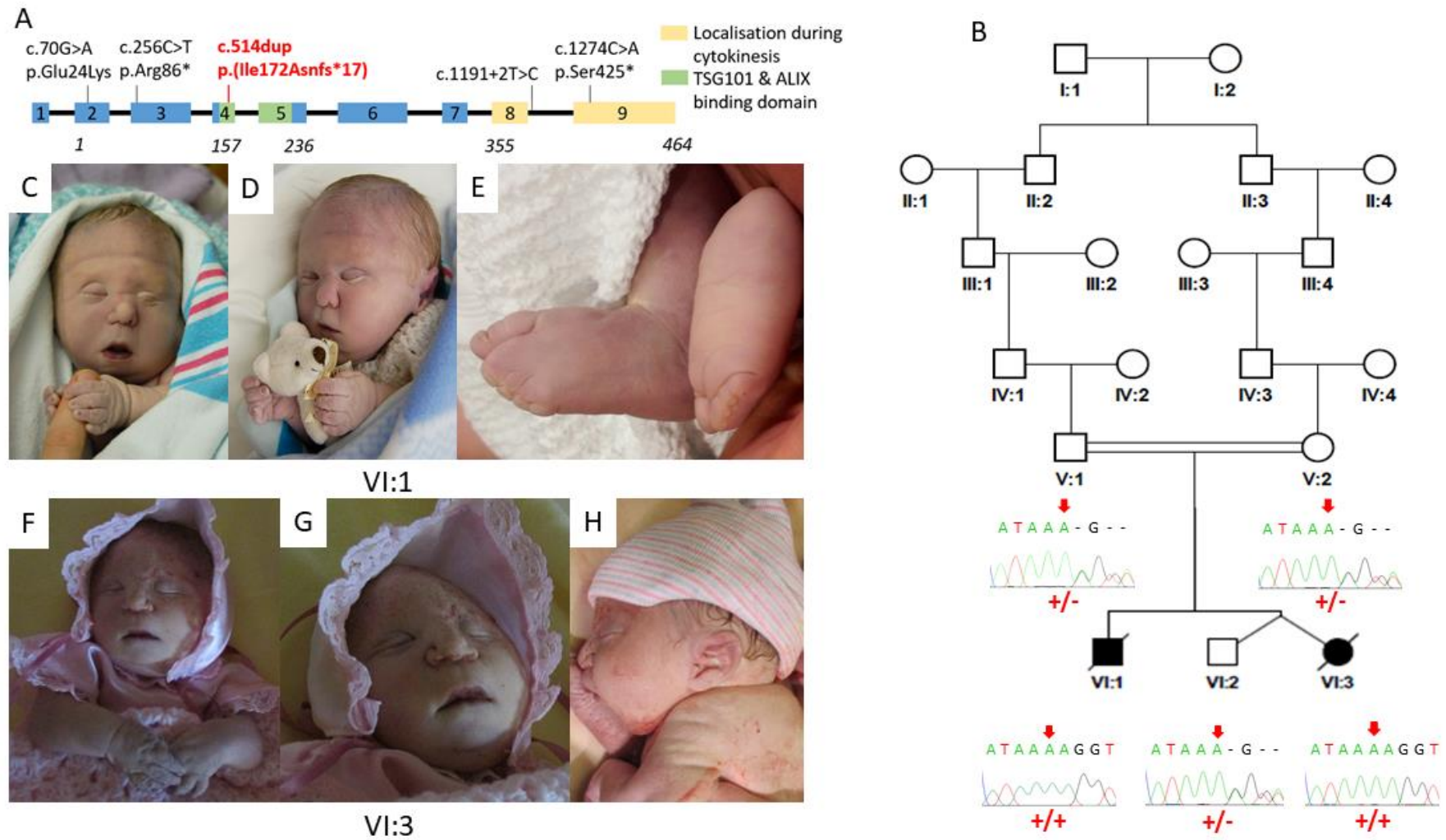
Figure 5.1: *CEP55* variants, family pedigree, genotype and images of affected individuals

(**A**) Schematic representation indicating the position of the disease-associated *CEP55* gene variants in relation to genomic organisation with the *CEP55* c.514dup; p.(Ile172Asnfs*17) variant identified in the current study highlighted by red text (exons; coloured boxes, introns; black lines). The correlation between coding regions and CEP55 polypeptide functional domains is also indicated; tumor suppressor gene 101 (TSG101) and apoptosis-linked gene 2 interacting protein X (ALIX) binding domains (green), and terminal region of the protein involved in localisation during cytokinesis (yellow). (**B**) Simplified pedigree of the Amish family investigated, with electropherograms showing the DNA sequence at the position of the *CEP55* variant (c.514dup) confirmed as homozygous in affected twin B and the deceased male sibling, and heterozygous in both parents and unaffected twin A. *CEP55* genotype is shown in red under electropherograms in generations V and VI (+, c.514dup; –, WT) (**C-H**) Clinical features of individuals homozygous for the *CEP55* variant (c.514dup). (**C-E**) VI:1 showing features of Potter sequence (also known as oligohydramnios sequence and used to describe a combination of distinctive facial and other associated phenotypic features that are a result of too little amniotic fluid. Features include epicanthic folds, retrognathia, a flattened nose, low-set ears, pulmonary hypoplasia and limb contractures including talipes) [441](15). Additional features observed included bilateral 2-5 toe syndactyly with a widened first web space and a bulbous nasal tip. (**F-H**) VI:3 (Twin B) showing features of Potter sequence. The features of both affected fetuses that can be attributed as secondary to oligohydramnios include redundant skin folds, short neck, flattened face, short pinched nose, retrognathia, small palpebral fissures and low set ears, brachydactyly, tapering fingers, and short 5th fingers with clinodactyly.

CASE	Bondesen *et al*	Bondesen *et al*	Frosk *et al*	Frosk *et al*	Frosk *et al*	Individual	Individual
Patient	II:2	II:3	303	305	306	VI:1	VI:3
CEP55 (NM_018131.4) Genotype	c.256C>T p.Arg86*	c.256C>T p.Arg86*	c.1274C>A p.S425*	c.1274C>A p.S425*	c.1274C>A p.S425*	c.514dup p.(Ile172Asnfs*17)	c.514dup p.(Ile172Asnfs*17)
Sex	M	M	M	M	M	M	F
Pregnancy outcome (gestation)	TOP (20)	IUFD (14+6)	SB (30)	ND (35)	SB (32)	SB (42)	ND (42)
Gestation anomalies identified	19 + 4	10 + 4	20	n/k	n/k	19	21 + 2
POTTER SEQUENCE	n/a	n/a	+	+	+	+	+
RENAL FEATURES							
Renal aplasia/dysplasia	+	n/k	+	+	+	n/k	+
Ureteral agenesis	n/k	n/k	+	+	+	n/k	n/k
Renal cysts	+	n/k	-	-	-	n/k	+
Bilateral hydronephrosis	+/-	n/k	-	-	-	+	-
Oligohydramnios	+	n/a	+	+	+	+	+
Pulmonary hypoplasia	n/k	n/k	+	+	+	n/k	n/k
Contractures	-	n/a	+	+	+	-	+
Talipes	+	n/a	+	+	+	+	+
CNS FEATURES							
Hydranencephaly	+	-	+	+	+	+	+
Cerebral cysts	-	+	-	-	-	-	-
Cerebellar hypoplasia	+	n/k	+	+	+	n/k	n/k
Encephalocele	-	+	-	-	-	-	-
Multinucleated neurons	n/k	n/k	+	+	+	n/k	n/k
Skull asymmetry	+	-	-	-	-	-	-
GROWTH							
IUGR	+	n/k	n/a	n/a	n/a	+	+
CARDIAC FEATURES							
Dysplastic aortic valve	n/k	n/k	+	n/k	+	n/k	n/k
Dilated left ventricle	n/k	n/k	+	n/k	+	n/k	n/k
Pericardial effusion	n/k	n/k	-	-	-	-	+
SKELETAL FEATURES							
Shortened/bowed long bones	n/k	n/a	-	-	-	+	+
Vertebral abnormalities	n/k	n/a	+	+	-	n/k	n/k
OTHER FEATURES							
Cystic hygroma	+	+	+	+	+	-	-
Redundant neck skin	n/a	n/a	+	+	+	+	+
Bulbous nasal tip	n/k	n/a	+	+	+	+	+
Syndactyly	n/a	n/a	+	+	+	+	+
Brachydactyly	n/a	n/a	+	+	+	+	+
Widened first web space	n/a	n/a	+	+	+	+	+
Single umbilical artery	+	n/a	-	-	-	-	+

Table 5.1: Comparison of cases with biallelic *CEP55* variants

(+), indicates presence of a feature in an affected subject; (-), indicates absence of a feature in an affected subject. Abbreviations: TOP, termination of pregnancy; IUFD, intrauterine fetal death; SB, stillbirth; ND, neonatal death; CNS, central nervous system; IUGR, intrauterine growth retardation; n/k, not known; n/a, not available.

Genetic studies

A combination of autozygosity mapping and WES were performed to study this syndrome and identify the underlying molecular cause, initially assuming that homozygosity for a founder gene variant was responsible for the condition although also considering other possible genetic causes. WES of DNA from the affected twin excluded previously described Amish founder variants associated with cystic kidney disease, including *NPHP3* [442]. Genome-wide SNP genotyping of DNA extracted from both twins identified a number of homozygous genomic regions particular to the affected individual, the largest of which entailed a ~20Mb region of 10q22.3-q24.1 (demarcated by rs1769756- rs7081796, chr10:g.79164647-99204526 [hg38]). Rare variants predicted to have a functional consequence were cross-referenced with SNP genotyping data, which identified a single stand-out candidate homozygous variant of relevance to the phenotype; a chr10:g.93507042dup, NM_018131.4: c.514dup; p.(Ile172Asnfs*17) [hg38] variant in *CEP55* located within the chromosome 10 locus, predicted to result in a premature stop codon. The variant is present in a single heterozygote in gnomAD v2.1.1/v3.1 and is not listed in ClinVar, NCBI or HGMDpro databases. Dideoxy sequencing confirmed the presence and cosegregation of the variant (Figure 5.1B). 179 samples from healthy Amish adults were analysed and seven heterozygous carriers were identified, corresponding to an estimated allele frequency of 0.02 in this population. Further, in a large database of >5000 Amish control individuals (Anabaptist Variant Server) the allele frequency for the p.(Ile172Asnfs*17) variant was 0.002465.

5.2.5 Discussion

This is the third reported family with likely homozygous loss of function variants in *CEP55* identified in association with a lethal fetal disorder (comparison of cases; Table 5.1). Frosk *et al* reported a family with Dutch-German Mennonite ancestry and three affected fetuses homozygous for a *CEP55* nonsense variant (NM_018131.4) c.1274C>A; p.Ser425* presenting with dysplastic kidneys hydranencephaly, cerebellar hypoplasia and multinucleated neurones at PM [437]. The authors termed this disorder MARCH syndrome (multinucleated neurons, anhydramnios, renal dysplasia, cerebellar hypoplasia and hydranencephaly syndrome [OMIM #236500]) and highlighted an additional nine cases in the literature with phenotypic overlap

including features of hydranencephaly, renal dysplasia and syndactyly [443-447]. However, as far as we are aware, *CEP55* genetic analysis has not been undertaken in these individuals. Interestingly in two cases neuropathological PM findings identified multinucleated neurons [443, 444]. Bondeson *et al* reported a Swedish couple with two affected fetuses homozygous for *CEP55* c.256C>T; p.Arg86* (NM_018131.4) with features in one including: hydranencephaly, enlarged cystic kidneys, oligohydramnios and cystic hygroma [438]. The second fetus had a slightly different phenotype comprising occipital encephalocele, cerebral cyst and cystic hygroma, the kidneys were severely degraded due to fetal autolysis. PM examination was only possible for one fetus and, although neither were identified to have polydactyly or liver abnormalities, the authors classified the combination of clinical features as a Meckel-like syndrome.

All seven cases described with biallelic *CEP55* loss of function variants display phenotypical similarity, with renal dysplasia or cysts resulting in oligohydramnios and Potter sequence, and CNS abnormalities comprising hydranencephaly or cerebral cysts, as the cardinal clinical features (Table 5.1). Interestingly, as noted by Bondeson *et al*, the presence of these two congenital anomalies fulfils two of the characteristic Meckel syndrome (MKS) clinical triad [438]. However, in the absence of polydactyly, none of the seven cases described to date fulfil the triad. Although PM examination was undertaken in only three cases, hepatic ductal plate malformation, a frequent finding in classical MKS, was not identified. Occipital encephalocele, the most frequent CNS abnormality in MKS (83.8% of cases) [448], was documented in one fetus [438]. A spectrum of other CNS abnormalities have been described in MKS cases, including cerebral cysts and hydrocephalus, of which hydranencephaly can be considered a severe form. Several other features reported in MKS were identified in this patient cohort, including syndactyly, hydronephrosis, and short bowed limbs. In view of the wide variability seen in MKS, the lack of diagnostic criteria, and the phenotypic overlap with CEP55 patients, this disorder should be considered a MKS-like condition [438].

The Amish siblings described here enable us to more precisely delineate the clinical consequences of CEP55 loss-of-function, defining hydranencephaly and cystic renal dysplasia as the predominant features. The identification of 21 additional Amish control samples heterozygous for the *CEP55* frameshift variant highlights the importance of testing when an Amish affected fetus presents with Potter sequence or

an MKS-like phenotype. Antenatal USS and PM investigations following stillbirth or neonatal death are infrequently undertaken by the Amish, and despite no previous reports, the allele frequency suggests that the condition is under recognised.

Unfortunately, while no tissue was available for cerebral histological analysis from the Amish and Bondeson cases [438], all three Frosk cases [437] displayed multinucleated neurons in cerebral tissue, and in one case multinucleated hepatocytes. Disruption of the CEP55 binding site for TSG101 and ALIX likely results from both the p.Arg86* and p.(Ile172Asnfs*17) variants (Figure 5.1A), which may prevent the cytokinesis abscission process resulting in incomplete cell division, providing a plausible explanation for the multinucleated neurons. The c.256C>T; p.Arg86* variant, located in exon 3 would be predicted to result in nonsense mediated mRNA decay, however studies on heterozygous parents identified equal levels of wild type and truncated transcript [438]. The p.Ser425* variant is predicted to delete the terminal 40 amino acids critical for localisation during cytokinesis, rather than lead to absence of the protein. Consistent with this, subcellular localisation studies showed the variant disrupts localisation to the midbody during cell division [437].

The lack of development of cerebral structures suggests that loss of CEP55 function may play a crucial role in cell division and migration during embryogenesis and development. Four recent studies of $Cep55^{-/-}$ mouse models demonstrate striking similarity with the human phenotype including perinatal lethality, polycystic kidneys, severe microcephaly and brain abnormalities including cerebellar hypoplasia, hydrocephalus, severe ventricular dilatation and cortical thinning/atrophy [435, 436, 439, 440]. $Cep55^{-/-}$ mice show a reduction in size of all brain areas, with reduced numbers of neurons in all cortical layers [436] and reduced numbers of neural stem/progenitor cells in the ventricular zone resulting from increased apoptosis [435]. In addition there were defects of abscission processes with an increased proportion of multinucleated neurons compared with controls [435, 440], consistent with the human phenotypes described. CEP55 has been identified to regulate ciliogenesis; $Cep55^{-/-}$ mice show a reduced number and shortening of cilia, with structural abnormalities and failure of cilia disassembly [435, 439]. Impaired cytokinesis may contribute to impaired neuronal migration in cells with aberrant cellular division, potentially explaining the MKS-like phenotype as both impaired cilia function and cell division processes may cause abnormal neuronal migration. However, further studies of CEP55 function are

required to fully determine the precise pathomolecular basis of this multisystem congenital anomaly disorder. Taken together, our findings consolidate CEP55 as a molecule fundamental to normal human development, with homozygous loss of function associated with a Meckel-like lethal fetal disorder profoundly affecting brain and kidney development.

Since the completion of this study, Barrie *et al* [449] described a non-lethal phenotype associated with biallelic *CEP55* variants identified in seven individuals from five unrelated families. The phenotype reported comprised of a neurodevelopmental disorder characterised by microcephaly (up to -7.1 standard deviations below the mean), developmental delay, toe syndactyly, dysmorphic features, behavioural abnormalities and brain abnormalities including white matter anomalies, lissencephaly/pachygyria in some individuals. Three female siblings were found to be homozygous for a consensus splice variant (NM_018131.4: c.1191+2T>C, p.?) within the penultimate exon and therefore likely to escape non-sense mediated decay (NMD) (Figure 5.1). The remaining four unrelated individuals were all compound heterozygous for the same missense variant (NM_018131.4: c.70G>A p.(Glu24Lys)) *in trans* with one of two nonsense variants (p.Arg86* or p.Ser425*) previously reported by Bondeson *et al* [438] (2 patients) and Frosk *et al* [437] (2 patients) respectively. No functional studies of these novel variants were included in this publication, which would be required to confirm their pathogenicity. Although, these findings suggest that loss of function variants in *CEP55* result in a severe lethal fetal disorder, whereas hypomorphic variants (c.1191+2T>C and p.(Glu24Lys)) with some residual CEP55 function may result in a milder phenotype. The c.1191+2T>C variant is predicted to escape NMD, as is the previously described p.Ser425* variant, located within the last exon of *CEP55*. Both variants are predicted to impact the C-terminus and likely result in disruption of protein localisation functions, yet the c.1191+2T>C variant is associated with a non-lethal phenotype. This milder phenotype may be explained by intron retention, partial exon skipping or 'leaky' splicing resulting from the c.1191+2T>C splice variant, rather than complete truncation of the protein. The recurrent variants identified in unrelated individuals (p.(Glu24Lys), p.Arg86* and p.Ser425*) may represent either mutational hotspots or founder mutations. Further studies of CEP55 function are required to fully determine the full phenotypic spectrum and precise pathomolecular basis of this multisystem congenital anomaly disorder.

5.3 Consolidating the clinical and genetic spectrum of inositol polyphosphate phosphatase *INPP4A*-related neurodevelopmental disorder

5.3.1 Summary

Inositol polyphosphate-4-phosphatase type IA (INPP4A) is a member of the highly conserved phosphoinositide phosphatase family of molecules that dephosphorylate phosphoinositides (PIs), ubiquitous membrane-associated signalling molecules involved in diverse cellular processes. Here clinical and genomic findings are described in eight unrelated families with 18 individuals affected by a severe neurodevelopmental disorder, in which genomic studies identified biallelic *INPP4A* (NM_001134225.2: c.137T>G; p.(Leu46*), c.115C>T; p.(Gln39*), c.352_353del; p.(Ser118Argfs*3), c.2278C>T; p.(Arg760Cys), c.322_323del; p.(Leu108Tyrfs*13), c.386C>A; p.(Thr129Lys) and c.2642G>A; p.(Arg881His)) gene variants as the likely cause of disease. Together with revisiting previously published data defining biallelic *INPP4A* gene variants as the candidate cause of disease in four families, this data consolidates disruption of INPP4A activity as the cause of a human neurodevelopmental disorder, and enables a more detailed description of the genetic basis and clinical features of the condition.

5.3.2 Introduction

Phosphatidylinositol (PtdIns) is a fundamental membrane phospholipid that can be reversibly phosphorylated at the 3, 4 and 5 positions of the inositol ring to produce phosphoinositides (PIs) at membrane bilayers [450]. PIs are important for a number of cellular processes including cytoskeleton remodelling, intracellular membrane trafficking, signal transduction, and control of cell growth and survival [450, 451]. The spatial and temporal expression of PIs are tightly regulated by specific phosphoinositide kinases and phosphatases [452]. Pathogenic variants within genes encoding phosphoinositide phosphatases have previously been associated with inherited neurodevelopmental disorders including *oculocerebrorenal syndrome of Lowe inositol polyphosphate-5-phosphatase* (*OCRL,* Lowe syndrome, OMIM #309000) [453], *inositol polyphosphate-5-phosphatase E* (*INPP5E,* Joubert syndrome 1/MORM (mental retardation, truncal obesity, congenital nonprogressive retinal dystrophy, and micropenis), OMIM #213300/#610156) [238, 454], *inositol*

polyphosphate-5-phosphatase K (*INPP5K*, congenital muscular dystrophy with intellectual disability and cataracts, OMIM #617404) [455] and *synaptojanin 1* (*SYNJ1*, developmental and epileptic encephalopathy 53, OMIM #617389) [456].

Inositol polyphosphate-4-phosphatase type IA (INPP4A) is an evolutionarily conserved phosphoinositide phosphatase that dephosphorylates phosphatidyl inositol-3,4-bis-phosphate (PtdIns(3,4)P2) to phosphatidyl inositol-3-phosphate (PtdIns(3)P) at the 4 position of the inositol ring [457]. PtdIns(3,4)P2 is an important membrane bound signalling molecule within the phosphatidylinositol-3-kinase (PI3K) pathway which activates protein kinase B (serine/threonine kinase 1, AKT) via 3-phosphoinositide-dependent kinase (PDK1) and mammalian target of rapamycin complex 2 (mTORC2) (Figure 5.2) [458, 459], thus acting as a negative downstream regulator of PI3K-AKT signalling. INPP4A (NM_001134225.2) is highly expressed in brain [460], has been shown to act as a suppressor of glutamate excitotoxicity regulating neuronal apoptosis within the CNS [461, 462].
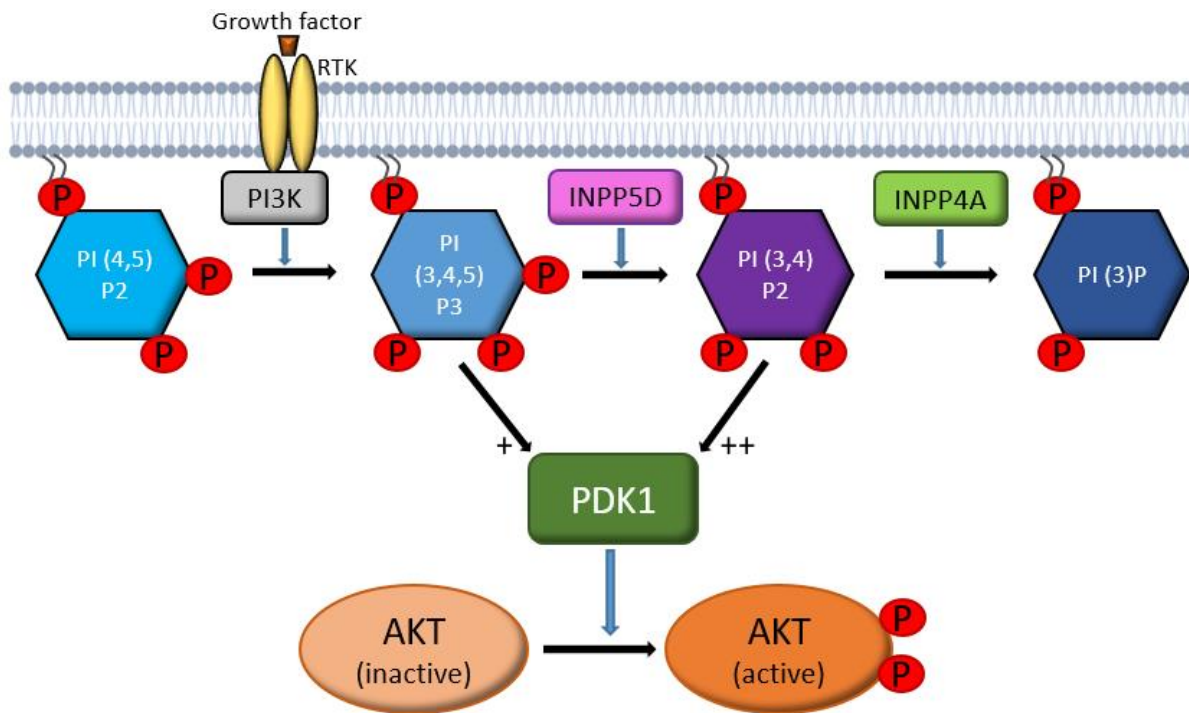
Figure 5.2: Schematic diagram of the proposed pathways involved in regulation of AKT activation by PDK1

Schematic figure depicting the possible role of Inositol polyphosphate-4-phosphatase type IA (INPP4A) in regulation of the PI3K-AKT pathway, through dephosphorylation of membrane bound phosphatidyl inositol-3,4-bis-phosphate (PtdIns(3,4)P2) to phosphatidyl inositol-3-phosphate (PtdIns(3)P) at the 4 position of the inositol ring. Loss of INPP4A leads to increased PtdIns(3,4)P2 accumulation, resulting in increased AKT activation via phosphorylation by PDK1. Abbreviations: RTK, receptor tyrosine kinase.

Previously, four publications described biallelic *INPP4A* variants as the candidate cause of disease in four families with a total of 11 individuals affected by a variable neurodevelopmental disorder with limited clinical information available and inconsistent phenotype shared by these families [463-466]. No functional studies were carried out to confirm pathogenicity of the identified variants. This study identifies eight further families with multiple affected individuals with an overlapping neurodevelopmental disorder in which biallelic *INPP4A* variants were identified as the likely cause of disease. Alongside reviewing and revisiting the genetic and clinical information in the previous studies, this work consolidates the genetic basis and more precisely defines the clinical spectrum of *INPP4A*-related neurodevelopmental disorder.

5.3.3 Methods

Families were recruited to the study with informed consent under approved protocols (see section 2.2.1). Blood samples were collected from all available family members and genomic DNA was extracted using standard techniques, as described in section 2.4.1. In Family 1, WES was performed in a single individual (IV:4) using Illumina NextSeq500 with Twist Human Core Exome targeting (University of Exeter). For Families 2-8 WES, variant prioritisation and analysis methodology were similar to those described in section 2.4.7. Primer design (Appendix 1), PCR and dideoxy sequencing confirmation of *INPP4A* variants was undertaken using standard techniques described in section 2.4.2. Clinical phenotyping was carried out according to methods detailed in section 2.2.2.

5.3.4 Results

Clinical and genetic studies

Table 5.2 details the clinical features of 12 affected individuals from eight unrelated families included in this study (Figure 5.2A), alongside updated clinical features of 11 individuals described in previously published studies (with an additional affected individual identified in Family 9) (Figure A5.1B). Table 5.3 summarises the clinical features of *INPP4A*-related disorder (See Appendix 5 section A5.1 for detailed clinical descriptions for each family). Affected individuals were identified through our established collaborative networks and GeneMatcher [244] in which WES studies identified rare (<0.01 frequency) biallelic *INPP4A* variants predicted to have a

deleterious impact on protein function (Figure 5.2B, Table 5.4). Dideoxy sequencing studies confirm appropriate segregation of *INPP4A* variants identified in Families 1, 2, 5, 6 and 9 (Appendix 5, Figure A5.1A), with dideoxy sequencing confirmation studies still ongoing for the remaining families. Appendix 5, Table 5.1 lists other variants identified by WES that could not be fully excluded for Families 1 and 5, with WES analysis ongoing for other families to exclude other possible candidate causes of disease.

FAMILY	Family 1: IV:4	Family 1: IV:5	Family 2: IV:4	Family 2: IV:6	Family 2: IV:7	Family 3: II:1
Genotype	c.137T>G; p.(Leu46*) Homozygous	c.137T>G; p.(Leu46*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.352_353del; p.(Ser118Argfs*3) Homozygous
Gender	Female	Female	Male	Male	Male	Male
Ethnicity	Pakistan (Pashtun)	Pakistan (Pashtun)	Iran (Arab)	Iran (Arab)	Iran (Arab)	Palestine (Arab)
Age at last examination (years)	25	21	40	34	38	27
Birth gestation (weeks)	Full-term	Full-term	NK	NK	NK	39
Antenatal/neonatal concerns	Delayed growth	Delayed growth	NK	NK	NK	Feeding problems
Growth parameters						
Birth weight kg (SD)	NK	NK	NK	5 (+2.85)	NK	3.45 (0.23)
Birth OFC cm (SD)	Normal	Normal	NK	NK	NK	Normal
Height cm (SD)	137.2 (-4.42)	111.8 (-8.59)	NK	NK	NK	160 (-2.44)
Weight kg (SD)	48 (-1.46)	37 (-3.69)	NK	NK	NK	39 (-4.61)
OFC cm (SD)	52 (-2.54)	45 (-7.58)	59 (+1.02)	58 (+0.43)	57 (-0.15)	52 (-3.07)
Development						
Intellectual disability	Severe	Severe	Severe (IQ 35-45)	Severe (IQ 35-45)	Severe (IQ 35-45)	Profound
Global developmental delay	Y	Y	Y	Y	Y	Y
Gross motor skills	Non-ambulatory Mobilises by crawling, sitting unsupported	Non-ambulatory Can only sit with support	Non-ambulatory Sitting unsupported	Non-ambulatory Sitting unsupported	Non-ambulatory Sitting unsupported, standing with support	Non-ambulatory Unable to sit unsupported
Speech delay	Severe, few words, partial understanding	Non-verbal, sounds only, very limited understanding	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words, partial understanding	Non-verbal, very limited understanding
Vision/eye abnormalities	Strabismus	Strabismus	N	N	N	Low visual acuity, normal VEP/ERG
Hearing problems	N	Uncertain, unresponsive	N	N	N	N
Neurology						
Any dysarthria	N	Non-verbal	Y	Y	Y	Non-verbal
Cerebellar signs	Nystagmus	Nystagmus	NK	NK	NK	Nystagmus
Gait abnormalities	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory
Central tone	NK	NK	Normal	Normal	Normal	Axial hypotonia
Muscle atrophy	Lower limbs	Lower limbs	Lower limbs	Lower limbs	Lower limbs	Upper and lower limbs, spastic quadriparesis

Upper limb	Mild upper limb weakness	Mild upper limb weakness	All normal	All normal	All normal	Hypertonia, weakness, hyperreflexia
Lower limb	Severe lower limb weakness	Severe lower limb weakness	Power 2/5, tone normal	Power 2/5, tone normal	Power 3-4/5, tone normal	Hypertonia, weakness, hyperreflexia
Plantar response	NK	NK	Flexor	Flexor	Flexor	Extensor
Seizures	N	N	N	N	N	Severe, intractable, onset 6 months, GTC, tonic, pharmacoresistant. AED: valproic acid, clonazepam
Neuroimaging findings	NP	NP	NP	NP	NP	CT: Cerebellar hypoplasia, cerebral atrophy, Dandy Walker variant, multiple hyperdensities
EEG findings	NP	NP	NP	NP	NP	NP
Behavioural features	Socially withdrawn, poor sleep	Socially withdrawn, avoids eye contact	N	N	N	Irritable
Other neurological features	N	Continuous head shaking (?titubation)	N	N	N	Choreoathetoid/dystonic movements, pyramidal signs, opisthotonus
Other features						
Dysmorphic features	Hypertelorism, curved toes	Hypertelorism, strabismus	Hypertelorism, happy expression	Rough face, happy expression	Happy expression	Coarse facial features, prominent ear lobes
Additional clinical findings	Dental anomalies, recurrent vomiting episodes	Dental anomalies, recurrent vomiting episodes	Dental anomalies, pitting of skin over the posterior scalp, raised BMI	Dental anomalies, pitting of skin over the posterior scalp, raised BMI	Dental anomalies	Gastrostomy, tracheostomy (room air), drooling, joint contractures
Other relevant clinical investigations	N	N	N	N	N	N

FAMILY	Family 4: II:1	Family 5: IV:2	Family 6: IV:1	Family 6: IV:2	Family 7: II:2	Family 8: II:1
Genotype	c.352_353del; p.(Ser118Argfs*3) Homozygous	c.322_323del; p.(Leu108Tyrfs*13) Homozygous	c.2278C>T; p.(Arg760Cys) Homozygous	c.2278C>T; p.(Arg760Cys) Homozygous	c.2642G>A; p.(Arg881His) Homozygous	c.386C>A; p.(Thr129Lys) Homozygous
Gender	Male	Female	Male	Female	Male	Female
Origin/Ethnicity	Palestine (Arab)	Egypt	Iran	Iran	Algeria (Kabylia)	Pakistan
Age at last examination (years)	7.5	13	10	7	21	6
Birth gestation (weeks)	40	40	38	38	39	39
Antenatal/neonatal concerns	Feeding problems	Nil	No	No	Nil	Nil
Growth parameters						
Birth weight kg (SD)	3.735 (0.37)	3 (-0.9)	3.1 (-0.06)	2.9 (-0.24)	2.98 (-0.79)	NK
Birth OFC cm (SD)	35 (-0.16)	34 (-0.44)	35 (-0.16)	34 (-0.44)	36 (+1.1)	NK
Height cm (SD)	120 (-0.92)	136 (-2.78)	128 (-1.68)	108 (-2.57)	175 (-0.3)	NK
Weight kg (SD)	18.5 (-2.23)	20 (-5.26)	25 (-1.57)	18 (-1.76)	93 (+2.18)	NK
OFC cm (SD)	48 (-3.65)	49 (-4.26)	54.5 (+0.01)	52.5 (-0.09)	62.5 (+3.06) father also has a large head	NK
Development						
Intellectual disability	Profound	Profound (IQ<20)	Moderate/severe	Moderate	Severe	Moderate
Global developmental delay	Y	Y (Regression at 1 year)	Y	Y	Y	Y
Gross motor skills	Non-ambulatory Sitting with support	Non-ambulatory Nil achieved	Walked aged 7, running, difficulty climbing stairs	Non-ambulatory Crawling	Walked aged 2	Walked age 2, left hemiparesis
Speech delay	Non-verbal, very limited understanding	Non-verbal, very limited understanding	Severe, 2 words only, partial understanding	Severe, 3 words only, partial understanding	Severe, few words, partial understanding	Severe, few words, partial understanding
Vision/eye abnormalities	Low visual acuity, abnormal ERG	Strabismus	N	Right esotropia	N	Strabismus, glaucoma
Hearing problems	N	N	N	N	N	N
Neurology						
Any dysarthria	Non-verbal	Non-verbal	Y, pseudobulbar	Y, pseudobulbar	N	Y
Cerebellar signs	Nystagmus	Nystagmus, intention tremor	Ataxia, tremor, nystagmus	Ataxia	Intention tremor, ataxia	Nystagmus
Gait abnormalities	Non-ambulatory	Non-ambulatory	Ataxia	Non-ambulatory	Ataxia	Unsteady gait
Central tone	Axial hypotonia	Axial hypotonia	Normal	Normal	Normal	Normal
Muscle atrophy	Upper and lower limbs, spastic quadriparesis	Severe spastic quadriparesis	N	N	N	Left hemiparesis
Upper limb	Hypertonia, weakness, hyperreflexia	Spasticity, brisk reflexes, weakness	Reflexes 3+	Reflexes 3+	Normal	Left hemiparesis

Lower limb	Hypertonia, weakness, hyperreflexia	Spasticity, brisk reflexes, weakness	Reflexes 3+	Reflexes AJ 2+, KJ 3+	Normal	Left hemiparesis
Plantar response	Extensor	Extensor	Flexor	Flexor	Equivocal	
Seizures	Severe, intractable, onset 6 months, GTC, tonic, pharmacoresistant. AED: Topiramate, clonazepam, leviteracitum	Severe, intractable, onset 1.5 years, GTC, myoclonic, tonic, encephalopathy. AED: Lamotrigine, sodium valproate, leviteracitum, clonazepam, topiramate	N	N	Severe, intractable, onset 11 years, GTC, absences, pharmacoresistant. AED: Lamotrigine, sodium valproate, clobazam	Severe, onset 2.5 years, GTC. AED: Sodium valproate
Neuroimaging findings	MRI: Cerebellar vermis hypoplasia, cerebral atrophy, dilated lateral ventricles and mega cisterna magna	MRI: Progressive vermian atrophy, mild cortical atrophy, left temporo-parietal cyst, dilated lateral ventricles, thin corpus callosum, thinning of brain stem	MRI: Cerebellar atrophy	MRI: Cerebellar atrophy	MRI: Vermian atrophy 15 years	NK
EEG findings	High voltage spike and wave in right parietal region	Bitemporal epileptogenic discharges	NP	NP	NK	NK
Behavioural features	Irritable	Irritability, agitation, poor sleep	Autistic features	Autistic features	Agitation, aggressive episodes, socially withdrawn	Aggressive episodes
Other neurological features	Choreoathetoid/dystonic movements, pyramidal signs, opisthotonus	Encephalopathy	N	N	Macrocephaly - present in the father (unaffected)	N
Other findings						
Dysmorphic features	Coarse facial features, prominent ear lobes	Long face, high forehead and anterior hairline, prominent nose, short philtrum, broad chin, low set ears	N	N	Large jaw, frontal bossing	Left sided ptosis, large tongue and lips, downslanting palpebral fissures, mid-face hypoplasia
Additional clinical findings	Gastrostomy, tracheostomy, drooling, joint contractures	Clenched hands and curved toes, acquired arthrogryposis, wrists, elbows, knees	Pes cavus	N	N	Drooling, joint hypermobility
Other relevant clinical investigations	Echo, array CGH, metabolic: normal	VEP: delayed latency	N	N	N	N

FAMILY	Family 9 Banihashemi *et al*: IV:4	Family 9 Banihashemi *et al*: IV:1	Family 9 Banihashemi *et al*: IV:2	Family 9 Banihashemi *et al*: IV:3	Family 9 Banihashemi *et al*: V:2	Family 9 Banihashemi *et al*: V:3
Genotype	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous	c.115C>T; p.(Gln39*) Homozygous
Gender	Female	Male	Male	Male	Male	Female
Origin/Ethnicity	Iran (Arab)	Iran (Arab)	Iran (Arab)	Iran (Arab)	Iran (Arab)	Iran (Arab)
Age at last examination (years)	33	32	30	23	12	4
Birth gestation (weeks)	NK	NK	NK	NK	NK	NK
Antenatal/neonatal concerns	NK	NK	NK	NK	NK	NK
Growth parameters						
Birth weight kg (SD)	NK	NK	NK	NK	NK	NK
Birth OFC cm (SD)	NK	NK	NK	NK	NK	NK
Height cm (SD)	NK	NK	NK	NK	NK	NK
Weight kg (SD)	NK	NK	NK	NK	NK	NK
OFC cm (SD)	54 (-1.09)	57 (-0.15)	57 (-0.15)	58 (+0.43)	NK	NK
Development						
Intellectual disability	Severe (IQ 35-45)	Severe (IQ 35-45)	Severe (IQ 35-45)	Severe (IQ 35-45)	Severe (IQ 35-45)	Severe (IQ 35-45)
Global developmental delay	Y	Y	Y	Y	Y	Y
Gross motor skills	Non-ambulatory Crawling	Non-ambulatory Crawling	Non-ambulatory, quadrupedal walking Crawling	Non-ambulatory Crawling	Non-ambulatory Crawling	Non-ambulatory Crawling
Speech delay	Severe, incomprehensible words	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words, partial understanding	Severe, incomprehensible words
Vision/eye abnormalities	Strabismus	N	Strabismus	N	Strabismus	N
Hearing problems	N	N	N	N	N	N
Neurology						
Any dysarthria	Y	Y	Y	Y	Y	Y
Cerebellar signs	NK	NK	NK	NK	NK	NK
Gait abnormalities	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory	Non-ambulatory
Central tone	Normal	Normal	Normal	Normal	Hypotonia at birth	NK
Muscle atrophy	Lower limbs	Lower limbs	Lower limbs	Lower limbs	Lower limbs	Lower limbs

Upper limb	Power 4/5	Power 4/5	Power 4/5	Power 4/5	Power 4/5	Power 4/5
Lower limb	Power 2/5	Power 2/5	Power 2/5	Power 2/5	Power 2/5	Power 2/5
Plantar response	Flexor	Flexor	Flexor	Flexor	Flexor	Flexor
Seizures	N	N	Non-specific, onset in infancy improved with age, no medication	N	N	N
Neuroimaging findings	MRI: Normal	MRI: Normal	MRI: Normal	MRI: Normal	MRI: Normal	NP
EEG findings	NP	NP	Polyspike epileptiform activity, generalized slow waves with no epileptiform discharge	NP	NP	NP
Behavioural features	NK	NK	NK	NK	NK	NK
Other neurological features	Head shaking	Head shaking	N	N	N	N
Other findings						
Dysmorphic features	Open mouth appearance, happy expression	Open mouth appearance, happy expression	Open mouth appearance, happy expression	Open mouth appearance, happy expression	Open mouth appearance, large low set ears, happy expression	NK
Additional clinical findings	Bilateral syndactyly feet	N	N	N	Camptodactyly all fingers, pes cavus, slim feet	N
Other relevant clinical investigations	N	N	N	N	N	N

FAMILY	Family 10 Olson *et al*: 26	Family 11 Sheffer *et al*: II-4	Family 11 Sheffer *et al*: II-1	Family 12 Najmabadi *et al*: 1	Family 12 Najmabadi *et al*: 2	Family 12 Najmabadi *et al*: 3
Genotype	c.2096G>A; p.(Arg699His) Homozygous	g.98555472_ 98557243del (1770bp deletion) Homozygous	g.98555472_ 98557243del (1770bp deletion) Homozygous	p.(Asp915fs)? Homozygous	p.(Asp915fs)? Homozygous	p.(Asp915fs)? Homozygous
Gender	Female	Male	Male	NK	NK	NK
Origin/Ethnicity	NK	Arab	Arab	Iran	Iran	Iran
Age at last examination (years)	13	1.5	0.5	NK	NK	NK
Birth gestation (weeks)	40	40	40	NK	NK	NK
Antenatal/neonatal concerns	Neonatal seizures	Poor suck	Deceased 6 months	NK	NK	NK
Growth parameters						
Birth weight kg (SD)	NK	3.5 (-0.11)	2.7	NK	NK	NK
Birth OFC cm (SD)	NK	33 (-1.73)	NK	NK	NK	NK
Height cm (SD)	NK	NK	NK	NK	NK	NK
Weight kg (SD)	NK	NK	NK	NK	NK	NK
OFC cm (SD)	Microcephaly	39.5 (-7.63)	Microcephaly	NK	NK	NK
Development						
Intellectual disability	Severe	Severe	Severe	Moderate	Moderate	Moderate
Global developmental delay	Y	Y	Y	Y	Y	Y
Gross motor skills	Non-ambulatory	No milestones	NK	NK	NK	NK
Speech delay	Y	Non-verbal	NA	Y	Y	Y
Vision/eye abnormalities	Cortical visual impairment	No fixing and following, cortical blindness	NK	NK	NK	NK
Hearing problems	NK	Normal ABR	NK	NK	NK	NK
Neurology						
Any dysarthria	NK	Non-verbal	NK	NK	NK	NK
Cerebellar signs	Tremor	Vertical nystagmus	NK	NK	NK	NK
Gait abnormalities	Non-ambulatory	Non-ambulatory	NK	NK	NK	NK
Central tone	NK	Severe axial hypotonia	NK	NK	NK	NK
Muscle atrophy	Spastic quadriparesis	Spasticity, fisted hands	NK	NK	NK	NK
Upper limb	Spastic quadriparesis	Spasticity	NK	NK	NK	NK
Lower	Spastic quadriparesis	Spasticity	NK	NK	NK	NK

Plantar response	NK	NK	NK	NK	NK	NK
Seizures	Neonatal onset: myoclonic, GTC, spasms, tonic, focal clonic, West syndrome	Myoclonic seizures - onset 8 months	NK	NK	NK	NK
Neuroimaging findings	MRI: Cortical atrophy, hypomyelination/dysmyelination, hypoplastic optic nerves	MRI: Hindbrain malformation, enlarged fourth/lateral ventricles, cerebellar hypoplasia	CT: Generalized brain atrophy with hypoplastic cerebellum and rudimentary vermis	NK	NK	NK
EEG findings	Burst suppression, hypsarrhythmia, multifocal sharp waves with abnormal slow background	Intermittent interictal polyspike epileptiform activity over centrotemporal regions, paucity of sleep features	NP	NK	NK	NK
Behavioural features	NK	Persistent agitation	NK	NK	NK	NK
Other neurological features	N	Myoclonic arm jerks	NK	NK	NK	NK
Other findings						
Dysmorphic features	NK	Prominent metopic suture, high arched palate, micrognathia	NK	NK	NK	NK
Additional clinical findings	Neurogenic scoliosis, hip dysplasia, PKU	Opisthotonus	NK	NK	NK	NK
Other clinical investigations	NK	N	NK	NK	NK	NK

Table 5.2: Clinical features of 12 families (24 individuals) affected by a neurodevelopmental disorder associated with biallelic variants in *INPP4A*

Shaded columns are previously published families. *Family 2 are distantly related to the family published by Banihashemi *et al* [466] (Family 9). (✓); indicates presence of a feature in an affected individual, (-); indicates absence of a feature in an affected individual. Height, weight and OFC Z-scores were calculated using a Microsoft Excel add-in to access growth references based on the LMS method [245] using UK1990 reference population [54]. Abbreviations: NK; not known, NP; not performed, SDS; standard deviation score, OFC; occipitofrontal circumference, GTC; generalised tonic-clonic, PKU; phenylketonuria.
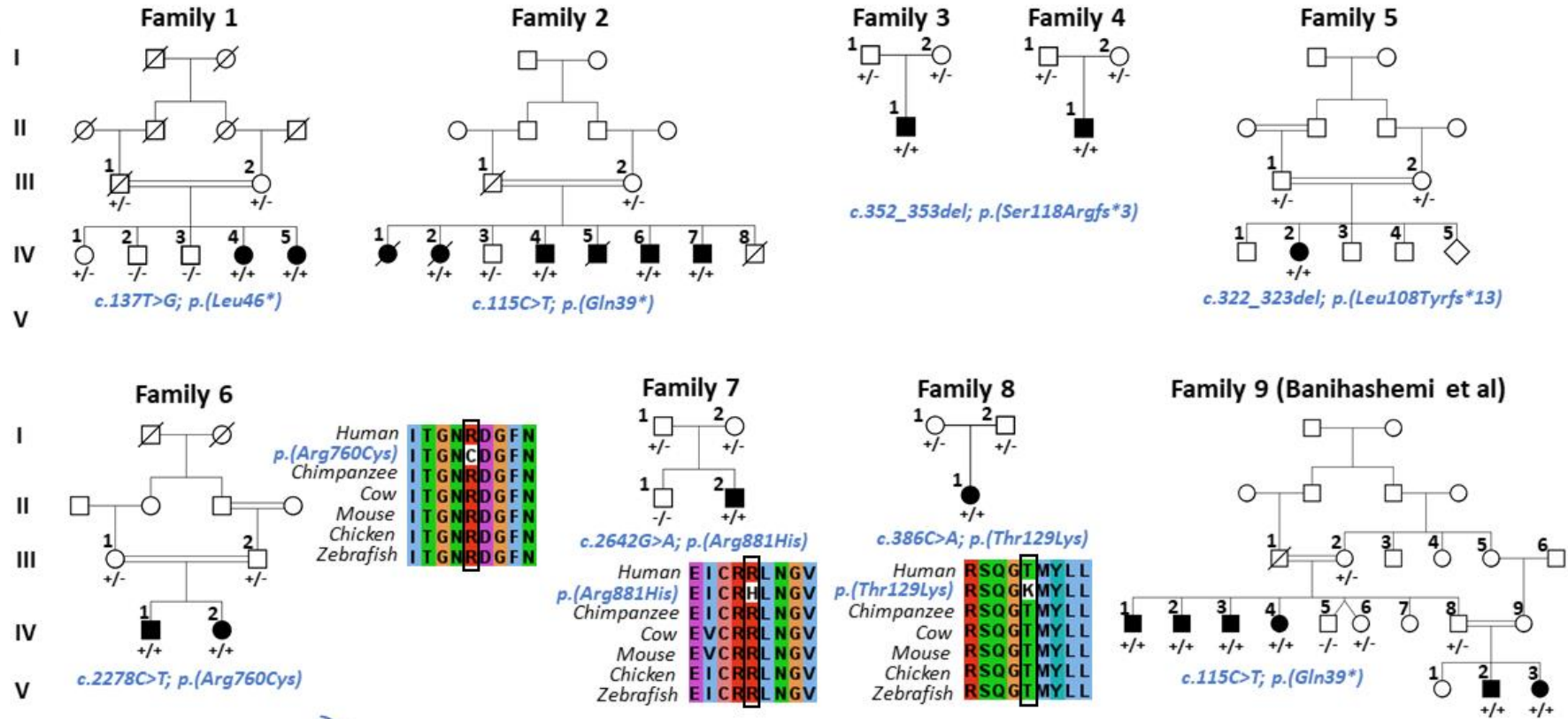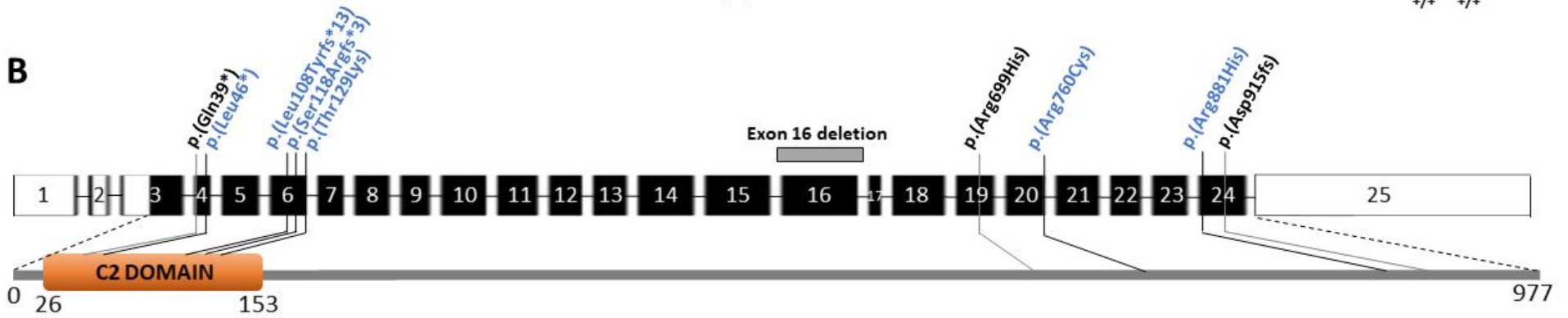
Figure 5.3: Schematic of *INPP4A* variants and family pedigrees

(**A**) Pedigrees of nine families (eight previously unreported (Families 1-8) and a previously described family (Family 9 (Banihashemi *et al* [466]), including an additional unpublished affected individual) with individuals affected by a neurodevelopmental disorder and biallelic variants within *INPP4A*. The presence of an *INPP4A* variant is indicated ('+'), with co-segregation confirmed in other family members ('-' indicates the wild type allele). Multi-species alignments are provided for missense variants (Families 5, 7 and 8) showing conservation of the molecular region encompassing each variant. (**B**) *INPP4A* gene schematic with solid black boxes indicate coding exons, white boxes indicate non-coding exons and below the protein structure displaying the location of a C2 domain and the identified *INPP4A* variants (blue variants are reported in the present study, black variants are previously described). A 1770bp deletion (Chr2:g.98555472_98557243del) including the whole of exon 16 is indicated by a grey box.

Taken together, our findings confirm the cardinal clinical features of *INPP4A*-related disorder as moderate to profound intellectual disability and global developmental delay in all affected individuals identified. This includes severe gross motor delay and speech impairment, with 84% of affected individuals non-ambulatory and 25% are non-verbal. Table 5.3 summarises the clinical features of *INPP4A*-related disorder in a total of 24 affected individuals. Cerebellar features of nystagmus, tremor, ataxia and speech difficulties were found in a large number of affected individuals, although this was a poorly documented clinical feature in many families. Other variable clinical features include lower limb weakness (80%), short stature (63%), microcephaly (44%), seizures (40%), spasticity (25%), striking dental anomalies including absent central incisors (25%), movement disorders including choreoathetoid, dystonic and myoclonic movements (15%), joint contracture (15%) and cortical visual impairment (10%). A behavioural phenotype includes features of agitation or irritability (46%), repetitive movements (40%) and autism (39%), with two individuals also displaying aggressive episodes. Neuroimaging (MRI/CT) was performed for 70% (14/20) of affected individuals with normal findings in 36% (5/14), cerebellar hypoplasia/atrophy in 57% (Figure 5.3E,F) and cortical atrophy in 43%. Other findings included ventricular dilatation in 21% (3/14) with megacisterna magna in 14% (2/14) and thinning of the corpus callosum in 14% (2/14) (Figure 5.3E,F). A Dandy Walker variant with hyperdense areas in bilateral parieto-occipital cortex and subcortical white matter was identified in Individual II:2 (Family 3), a large left temporo-parietal cyst in Individual II:2 from Family 5 (Figure 5.3G,H) and hypomyelination/dysmyelination with hypoplastic optic nerves in Patient 26 from Family 10 (Olson *et al* [465]). EEG studies were performed for five affected individuals and included findings of interictal polyspike epileptiform activity, generalized slow waves, hypsarrythmia and burst suppression (See Appendix 5 A5.1 for detailed EEG findings).

Clinical photographs and videos were reviewed for Families 1, 2, 5, 6 and 9, although consent for publication has currently only been obtained for Family 1 with facial features of hypertelorism, strabismus and absent central incisors (Figure 5.3A-D) observed in two affected individuals. These findings are strikingly similar to those in affected individuals from Family 2 and permission for publication of photographs for these families is being sought.

Clinical Feature	Biallelic LOF (19)	Biallelic missense (5)	Total cases (24)
Sex	F = 5, M = 11	F = 3, M = 2	F = 8, M = 13
Mean age in years (range)	22.0 (0.5 – 40)	11.4 (6 – 21)	19.5 (0.5 – 40)
Developmental phenotype			
Intellectual disability	100% (19/19)	100% (5/5)	100% (24/24)
- Moderate	16% (3/19)	60% (3/5)	25% (6/24)
- Severe	68% (13/19)	40% (2/5)	62.5% (15/24)
- Profound	16% (3/19)	0% (0/5)	12.5% (3/24)
Global developmental delay	100% (19/19)	100% (5/5)	100% (24/24)
Speech delay	100% (18/18)	100% (5/5)	100% (23/23)
- Non-verbal	33% (5/15)	0% (0/5)	11% (5/20)
Gross motor delay	100% (17/17)	100% (5/5)	100% (22/22)
- Non-ambulatory	100% (14/14)	40% (2/5)	84% (16/19)
Neurological features			
Microcephaly (≥2SD)	50% (7/14)	25% (1/4)	44% (8/18)
Lower limb weakness	100% (15/15)	20% (1/5)	80% (16/20)
Cerebellar features	100% (6/6)	100% (5/5)	100% (11/11)
Seizures	33% (5/15)	60% (3/5)	40% (8/20)
- Severe, intractable	60% (3/5)	100% (3/3)	75% (6/8)
- Generalised tonic-clonic	60% (3/5)	100% (3/3)	75% (6/8)
- Tonic	80% (4/5)	0% (0/3)	50% (4/8)
- Myoclonic	60% (3/5)	0% (0/3)	38% (3/8)
Spasticity	27% (4/15)	20% (1/5)	25% (5/20)
Movement disorder	20% (3/15)	0% (0/5)	15% (3/20)
Neuroimaging	63% (10/16)	100% (4/4)	70% (14/20)
- Cerebellar hypoplasia/atrophy	50% (5/10)	75% (3/4)	57% (8/14)
- Cortical atrophy	50% (5/10)	25% (1/4)	43% (6/14)
Behavioural phenotype			
Agitation/irritability	44% (4/9)	50% (2/4)	46% (6/13)
Repetitive movements	55% (6/11)	0% (0/4)	40% (6/15)
Impaired social interaction	22% (2/9)	75% (3/4)	39% (5/13)
Other features			
Short stature	80% (4/5)	33% (1/3)	63% (5/8)
Dental anomalies	33% (5/15)	0% (0/5)	25% (5/20)
Joint contractures/arthrogryposis	20% (3/15)	0% (0/5)	15% (3/20)

Table 5.3: Summary of clinical features of 24 individuals with biallelic *INPP4A* variants

Percentage of affected individuals with each clinical feature broken down into loss of function (LOF) and missense variants. Abbreviations: M; male, F; female, SD; standard deviation.

WES studies identified biallelic loss of function *INPP4A* variants in 19 affected individuals from Families 1-5, 9, 11 and 12 (see Table 5.4 for a list of variants). Table 5.3 provides a summary of the clinical features in affected individuals with *INPP4A*-related disorder divided into biallelic LOF or missense variant. Affected individuals with biallelic LOF variants generally have a more severe developmental phenotype, with all these individuals being non-ambulatory and 33% are non-verbal.

Biallelic missense variants were identified in five affected individuals from four families (Families 6-8 and 10). Although this is a small number of individuals, generally there appears to be a milder degree of intellectual disability and developmental delay (Table 5.3) including no non-verbal individuals, although a greater proportion of these individuals appear to have seizures and autistic features. Only two identified missense variants p.(Arg760Cys) and p.(Arg881His) are present in online databases (gnomAD v2.1.1) which lists one heterozygous (1 Finnish European, AF 0.000004056) and two heterozygous (1 South Asian, 1 non-Finnish European, AF 0.000008067) respectively and no homozygous individuals (Table 5.3). The p.(Arg699His) variant previously published by Olson *et al* [465] is present at low frequency in gnomAD v2.1.1 (7 heterozygotes) and v3.1.2 (5 heterozygotes) and was classified as a VUS (Table 5.3). None of these variants are listed in ClinVar (NCBI) or Human Gene Mutation Database (HGMD) Professional. Identified missense variants p.(Arg760Cys), p.(Arg881His), p.(Thr129Lys) and p.(Arg699His) are located within highly conserved regions, with stringent amino acid conservation to at least *Danio rerio* (Figures 5.2A and A5.1B), and are predicted to be pathogenic by multiple prediction software tools including Provean, SIFT, Polyphen-2 and REVEL (Table 5.3). Overall, only 18 loss of function variants are listed in the *INPP4A* gene in gnomAD (v2.1.1/v3.1.2) all at low allele frequency (< 0.00002485) and in heterozygous state only. There is a significant degree of constraint for *INPP4A* in gnomAD v2.1.1 for both pLOF (probability of being loss-of-function intolerant (pLI) score 1.0, observed/expected LoF=0.13, LOEUF 0.25) and missense variants (Z-score 3.35).

CASES	*INPP4A* NM_001134225.2	GRCh38 (hg38):	gnomAD v2.1.1 HET	gnomAD v3.1.2 HET	gnomAD HOM	gnomAD All AF	SIFT	Provean	Polyphen-2	REVEL	Splice AI
Published here:											
Family 1	c.137T>G; p.(Leu46*)	Chr2:g.98520717T>G	0	0	0	0	NA				
Families 3 and 4	c.352_353del; p.(Ser118Argfs*3)	Chr2:g.98535810_98535811del	0	0	0	0	NA				
Family 5	c.322_323del; p.(Leu108Tyrfs*13)	Chr2:g.98535780_98535781del	0	0	0	0	NA				
Family 6	c.2278C>T; p.(Arg760Cys)	Chr2:g.98565765C>T	1	0	0	0.000004056	0.001	-6.14	0.999	0.662	0.02
Family 7	c.2642G>A; p.(Arg881His)	Chr2:g.98576999G>A	2	0	0	0.000008067	0.002	-3.62	0.999	0.568	0.00
Family 8	c.386C>A; p.(Thr129Lys)	Chr2:g.98535844C>A	0	0	0	0	0.001	-3.84	0.999	0.356	0.05
Previously published:											
Families 2 and 9 Banihashemi et al	c.115C>T; p.(Gln39*)	Chr2:g.98520695C>T	0	0	0	0	NA				
Family 10 Olson et al	c.2096G>A; p.(Arg699His)	Chr2:g.98564707G>A	7	5	0	0.00003331	0.072	-1.52	0.904	0.209	0.00
Family 11 Sheffer et al	c.1566del256; p.(Glu523Ilefs*22)	Chr2:g.98555472_98557243del (1770bp deletion)	0	0	0	0	NA				
Family 12 Najmabadi et al	c.?; p.(Asp915fs)	NK	0	0	0	0	NA				

Table 5.4: Homozygous *INPP4A* (NM_001134225.2) variants identified in affected individuals from 12 families

Abbreviations: HOM; homozygous, AF; allele frequency, NA; not applicable, NK; not known.
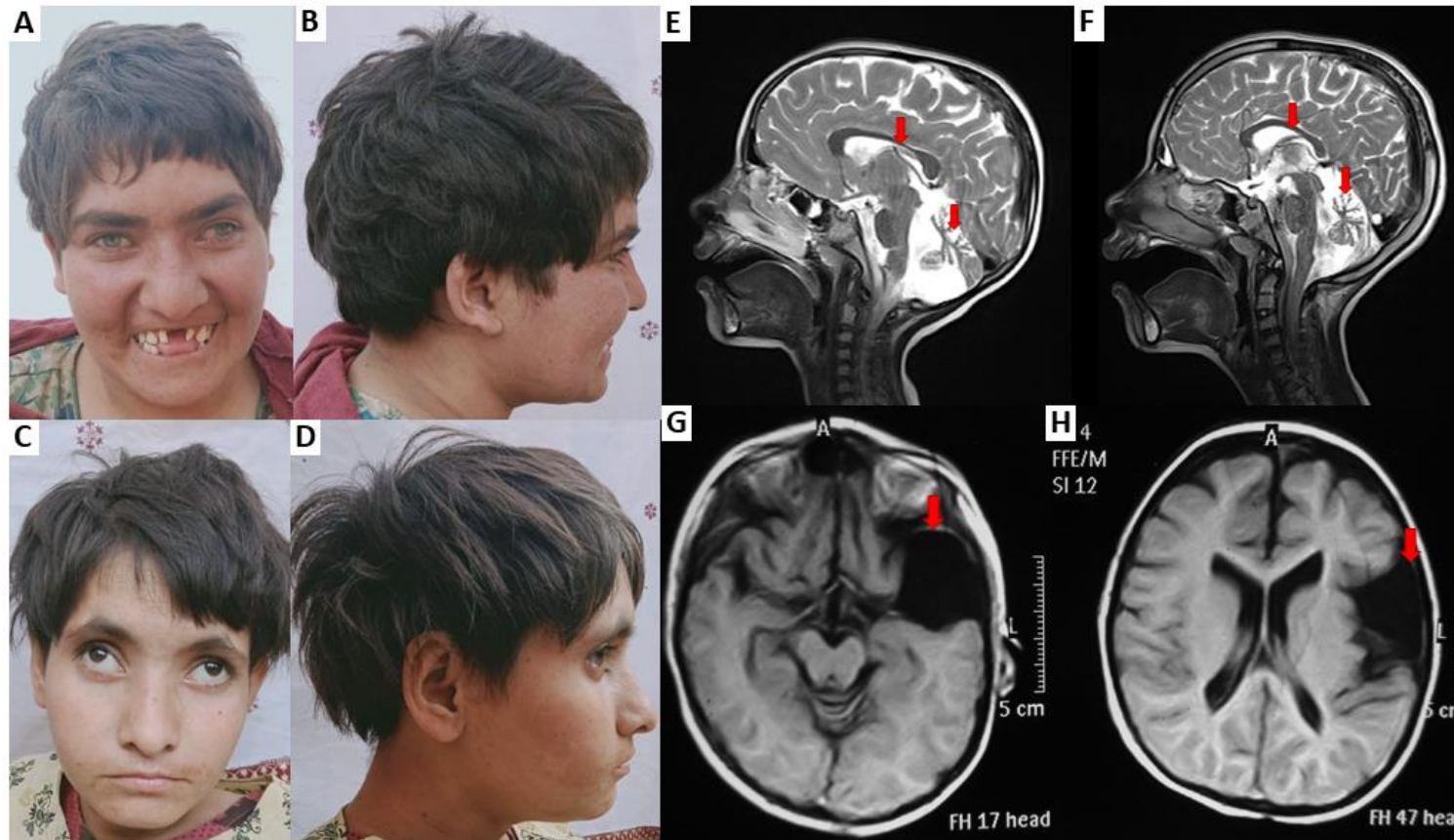
Figure 5.4: Clinical photographs and MRI brain imaging of individuals with *INPP4A*-related disorder

(**A-D**) Clinical photographs of two sisters (**A-B** of Individual IV:4 and **C-D** of Individual IV:5) from Family, 1 homozygous for the NM_001134225.2: c.137T>G; p.(Leu46*) variant in *INPP4A* with features of hypertelorism, strabismus and absent central incisors. (**E-H**) MRI brain imaging of three individuals with *INPP4A*-related disorder, two siblings from Family 6: Individual IV:2 (**E**) and Individual IV:1 (**F**) both with features of cerebellar hypoplasia and thinning of the corpus callosum, and Individual IV:2, Family 5 (**G-H**) showing a large left temporo-parietal cyst, mild cortical atrophy and mild dilatation of the lateral ventricles.

5.3.5 Discussion

With the description of eight new families (alongside reassessment of a previously reported family [466]; Family 9) described here, a total of 12 families (27 individuals) have now been described with likely *INPP4A*-related neurodevelopmental disorder (Table 5.2). The condition in eight families (22 individuals) is associated with biallelic protein truncating variants (two nonsense, three frameshift and one homozygous 1770bp deletion encompassing exon 15 and leading to a frameshift), and in four families (5 individuals) is associated with biallelic suspected pathogenic missense variants (Figure 5.2A, Table 5.3).

The core clinical features observed in all individuals with *INPP4A*-related disorder include global developmental delay with moderate to profound intellectual disability, severe speech impairment and cerebellar signs including nystagmus and unsteady gait. Other variable features include microcephaly, short stature, delayed motor development, muscle weakness, hypotonia, spasticity, seizures, involuntary movements and behavioural abnormalities. The age of onset of the condition is in infancy and between birth and 4 years of age. Microcephaly is reported in 44% (8/18) of affected individuals, although OFC measurements were not available for several affected individuals and not reported in several previously published patients. A single individual (Family 7) was reported with macrocephaly, although this may be a familial trait observed also in the father. All the affected individuals from families 1, 3, 5 and 6 reported here display significant short stature (-2.57 and -8.59SD below the mean), however this has not been noted in affected individuals from the other families. However, as height measurements are not available for all previously published patients it is unclear how consistent a feature of the disorder this is. Both receptive and expressive language appear to be universally impaired with affected individuals speaking only a few words or being non-verbal, typically expressive language appears to be more significantly impaired than receptive for most individuals. Motor features are highly variable from mild delay in attainment of motor milestones in six affected individuals, 18 affected individuals are non-ambulatory and achieve mobility through crawling or quadrupedal walking and three severely affected individuals are unable to sit without support. Lower limb weakness appears to be the cause of non-ambulation in most individuals, with a further five individuals reported to have spastic quadriparesis. Seizures are described in eight individuals with *INPP4A*-related

disorder and include generalised tonic-clonic, absence, myotonic, tonic and focal clonic seizure types reported with age of onset between the neonatal period to 11 years of age, with pharmacoresistance reported in four individuals. Five individuals underwent EEG investigation revealing features of generalised slowing and interictal polyspike epileptiform activity. All affected individuals have cerebellar features including nystagmus, strabismus, intention tremor and wide based/unsteady gait in those who are ambulatory. Six individuals with MRI brain imaging have cerebellar hypoplasia and vermian atrophy consistent with these findings. Other MRI brain findings noted include diffuse cerebral atrophy, hindbrain malformation, enlarged 4th and lateral ventricles, hypo/dysmyelination and hypoplastic optic nerves in a single individual. Five individuals from Family 9 had no abnormalities on MRI brain imaging. Additional ophthalmic features include cortical visual impairment in two individuals (one with hypoplastic optic nerves on MRI), reduced visual acuity and glaucoma. Behavioural features have been observed in several individuals and include impaired social interaction, head shaking, sleep disturbance, agitation, irritability and aggression. Several individuals have dental anomalies with absent central teeth. Other reported features are summarised in Tables 5.2 and 5.3.

In general, it appears that individuals with biallelic LOF variants display more severe disease than those harbouring missense alterations (Tables 5.2 and 5.3); all nineteen individuals from Families 1-5, 9 (Banihashemi *et al* [466]) and 11 (Sheffer *et al* [464]) display severe/profound intellectual disability, are non-ambulatory and non-verbal (except for Individual IV:4 in Family 1 who has a few words only). This may be explained by these protein truncating variants undergoing NMD with no resultant polypeptide product generated from the mutant allele [467, 468], leading to a more severe neurodevelopmental phenotype than that in individuals harbouring missense alterations in which some protein functionality may be preserved. Interestingly, there appears to be distinct phenotypes associated with nonsense and frameshift *INPP4A* variants. Families 1, 2 and 9 (Banihashemi *et al* [466]) harbour homozygous nonsense variants (p.(Gln39*) and p.(Leu46*)) and have strikingly similar phenotypes (Table 5.2) including severe intellectual disability, developmental delay, predominantly lower limb weakness without spasticity, cerebellar features and are non-ambulatory with quadrupedal walking, an absence of seizures and several individuals are identified to have dental abnormalities and head shaking. Further investigation is required as

several individuals were reported to have normal MRI neuroimaging, although in the majority this was not performed. Families 3, 4, 5 and 11 (Sheffer *et al* [464]) all have frameshift variants again with overlapping distinct phenotypes including severe/profound intellectual disability, severe spasticity with opisthotonus and severe intractable seizures, with several individuals displaying movement disorders. Individuals in these families had abnormalities on MRI brain that included cerebellar hypoplasia, cerebral atrophy, dilated asymmetric lateral ventricles and thin corpus callosum; with Dandy-Walker variant and left temporo-parietal cyst observed in single individuals.

To date five individuals with candidate biallelic missense *INPP4A* variants have been identified (Table 5.2) who in general display a milder neurodevelopmental phenotype with two walking at age 2 years and one at 7 years. Three missense variants were discovered in Families 6-8 of Iranian, Algerian and Pakistani origin and it is worth considering that these groups are underrepresented in available genomic databases, for example gnomAD, and it can often be difficult to exclude missense variants in these groups. Indeed, there are some phenotypic discrepancies in these individuals, for example Individual II:1 (Family 7) who has macrocephaly. All identified missense variants are classified as VUS according to ACMG and UK ACGS criteria [401, 402]. Further functional studies are currently underway in collaboration with Prof Volker Haucke's research group (Leibniz Research Institute for Molecular Pharmacology, Berlin) to examine the molecular effects of these four missense variants on INPP4A protein function to confirm their pathogenicity and to explore how this translates to observed differences in clinical phenotype.

Seizures appear to be of multiple types, often intractable and do not appear to correlate with the severity of cognitive impairment (Table 5.2), indicating a generally reduced seizure threshold in *INPP4A*-related disorder patients. Generalised spasticity is a feature observed in a subset of six patients with *INPP4A*-related disorder and appears to be associated with the presence of severe intractable seizures (Table 5.2); movement disorders (including choreoathetosis, opisthotonus, dystonia and myoclonic jerks) and joint contractures were additionally observed in three of these individuals. Interestingly individuals from two further families without features of spasticity (Families 1 and 9) were noted to display head shaking behaviours which could represent dystonia or tic-like disorder.

Two homozygous *INPP4A* knockout mouse models have been reported including the *weeble* mouse that has a spontaneous homozygous mutation in *Inpp4a* with a single nucleotide deletion leading to a frameshift [460] and a targeted *Inpp4A*$^{-/-}$ knockout mouse strain [462]. Both mouse models display an overlapping phenotype with that seen in *INPP4A*-related neurodevelopmental disorder in humans involving severe movement disorder being unable to walk, with poor weight gain and death at 3-4 weeks from severe seizures with rigidity. However, there are subtle phenotypical differences present in each mouse model, with the *weeble* mouse being ataxic and the *Inpp4A*$^{-/-}$ mouse displaying an involuntary movement disorder including limb hyperkinesias (choreas, ballism), opisthotonus and dystonia. Additionally, there are differences in the neuroanatomical defects present in each mouse consistent with their differing motor phenotypes, with the *weeble* mouse displaying a severe defect of the cerebellum and hippocampus with pronounced Purkinje and pyramidal cell degeneration associated with cell death during postnatal development and milder cell loss generally within the cerebral cortex. Histological examination of the *Inpp4A*$^{-/-}$ mouse identified degeneration of the striatum with medium-sized spiny projection neurons particularly affected. Other overlapping phenotypical features between humans and mouse models include reduced growth parameters, severe seizures and motor manifestations including inability to walk, cerebellar features and involuntary movements including opisthotonus, chorea, dystonia and myoclonic jerks. Interestingly several patients have neuroanatomical similarities to the mouse models on brain MRI imaging with cerebellar hypoplasia and generalised atrophy, although no individuals were identified to have basal ganglia degeneration.

INPP4A catalyses the dephosphorylation of PtdIns(3,4)P2 to PtdIns(3)P, therefore absence of INPP4A leads to delayed degradation of PtdIns(3,4)P2 and increased downstream PI3K-AKT signalling (Figure 5.2) [458, 459]. Other well characterised neurodevelopmental conditions associated with impaired PI3K-AKT signalling resulting in downstream overactivation of mTORC1 [166] are associated with overgrowth, including developmental brain abnormalities. One mechanism of pathogenicity of loss of INPP4A could be through increased PI3K-AKT signalling leading to reduced mTORC1 activation and resulting in poor growth, small stature and microcephaly. Consistent with this, INPP4A has been identified to play a key role in cell cycle regulation and apoptosis, and shuttles between the cytoplasm and nucleus

to achieve this [469]. INPP4A plays a protective role against NMDA receptor-mediated excitotoxic death of neurons, through regulation of NMDAR localisation at the post-synaptic density [462]. Lack of INPP4A expression in *Inpp4A*$^{-/-}$ mouse striatal cells leads to an increase in post synaptic NMDARs with larger NMDA-mediated currents observed, and potentiated glutamate-induced striatal neuron cell death, with toxic effects of glutamate at lower concentrations. Further in vitro studies confirm that knockdown of INPP4A in an epilepsy cell model results in increased apoptosis through increased calcium release [461]. This explains the increased susceptibility to seizure development and neurodegeneration in both *Inpp4A*$^{-/-}$ mice and individuals with *INPP4A*-related disorder. Further studies have identified that expression of a specific glutamate transporter (SLC1A6) may have a protective effect on Purkinje neuronal cell loss observed with absence of INPP4A in weeble mice [470]. Expression levels of other proteins affecting neuronal excitotoxicity in different regions of the brain may explain the severity and variability of features observed in individuals with *INPP4A*-related disorder, particularly seizures.

Here we notably extend the number of families and individuals affected by *INPP4A*-related neurodevelopmental disorder, enabling a more precise characterisation of the genetic basis and clinical phenotype of the condition. Importantly, previous studies in *Inpp4A*$^{-/-}$ mice show that NMDAR antagonist MK-801 improves hindlimb clasping and reduces seizures [462] suggesting that pharmacotherapeutic developments may well be possible for treatment of *INPP4A*-related neurodevelopmental disorder patients. Given that the age of onset of the condition is typically within the first year of life following an initial period of apparently normal development, there may be a potential therapeutic window for early intervention using NMDAR antagonists to alleviate certain aspects of the condition which should be explored in future research studies.

5.4 Conclusions and future work

Studies of Amish and Pakistani families with neurodevelopmental disorders, facilitating collaboration with other groups internationally, have aided confirmation of disease-gene associations for genetic disorders associated with variants in two genes: *CEP55* and *INPP4A*.

Interestingly, since our studies of *CEP55* were published a further paper has reported variants in *CEP55* associated with a non-lethal developmental disorder characterised by severe microcephaly, cognitive impairment, brain abnormalities (white matter defects and lissencephaly), dysmorphic features and behavioural problems in seven individuals [449]. Four unrelated individuals were compound heterozygous for a previously published protein truncating variant alongside a missense variant, and three siblings were homozygous for a canonical splice variant, likely acting as hypomorphic variants with some residual protein function. This identifies phenotypic and genotypic heterogeneity of *CEP55*-associated disorder, indicating that there may be genotype-phenotype correlations, a finding observed in many neurodevelopmental disorders. This is an important consideration when investigating established disease genes that milder disease may be associated with less deleterious gene variants that expand the associated clinical phenotype. In future, further studies could attempt to unravel the molecular basis for these differing phenotypes. This may involve defining the precise roles of the CEP55 interacting domains and establishing the effect of protein truncating variants on transcript and protein stability, binding of interacting partners, downstream signalling pathways and cilia function, for example, through the generation of variant constructs and patient cell lines. Structural modelling of missense variants and splicing studies to determine the outcome of splice variants should also be considered. Additionally, the identification of further individuals with *CEP55* associated disorder will help further characterisation of the genetic and phenotypic spectrum and the establishment of clinical management guidelines.

This study also consolidates biallelic variants in *INPP4A* as causative of an autosomal recessive neurodevelopmental disorder frequently associated with microcephaly. Further clinical studies in additional individuals affected with *INPP4A*-associated disorder subsequently identified will enable further clarification of the full clinical spectrum of disease, particularly aiding our understanding of the basis of the presence

or lack of generalised spasticity, and confirmation of whether dental anomalies are an associated feature. INPP4A is a phosphoinositide phosphatase that dephosphorylates PtdIns(3,4)P2, a negative regulator of the downstream PI3K-AKT signalling pathway involved in many cellular processes [457-459]. Functional studies, which are currently lacking, will provide support for the pathogenicity of the *INPP4A* variants identified in individuals with this disorder, provide insights into the pathomechanistic basis of disease and potentially enable genotype-phenotype correlations to be drawn. Such studies may include analysis of patient cells (for example lymphoblastoid cell lines) to investigate the impact of gene variants on protein levels, and also phosphatase activity on PtdIns(3,4)P2 and downstream effects on the PI3K-AKT pathway. Prof Volker Haucke's research group (Leibniz Research Institute for Molecular Pharmacology, Berlin) with expertise in cellular endocytic and endolysosomal systems are currently undertaking functional studies to support the pathogenicity of the missense variants identified in this study, for inclusion in our pending manuscript detailing our data. Despite the current lack of functional studies, the number of families (including several large families) with segregating rare likely deleterious variants associated with the close phenotypical overlap in affected individuals provides strong evidence for the gene disease association. As such *INPP4A* should be considered for inclusion in gene panels for intellectual disability. In *Inpp4A*$^{-/-}$ mice the absence of INPP4A has been associated with NMDA receptor-mediated excitotoxic death of neurons and an increased susceptibility to seizures [462]. As the NMDA receptor antagonist MK-801 improved clinical symptoms in mice [462] it is important to undertake studies of potential future treatments of *INPP4A*-associated neurodevelopmental disorder in humans. This is of particular relevance as there may be a brief period of normal developmental before the onset of symptoms (between 5 months and 4 years) that could be a potential therapeutic window to prevent the development of some features of this condition.

6 CONCLUDING COMMENTS

This thesis documents the identification and/or consolidation of four genetic disorders arising due to sequence variants in the *TRAPPC10*, *KPTN*, *CEP55 and INPP4A* genes, each causative of a rare autosomal recessive disorder of brain growth and development. Each condition was identified as part of well-established genetic research programmes aiming to discover the spectrum and nature of genetic disorders present in Amish and Pakistani communities, affecting families with undiagnosed neurodevelopmental disorders. These programmes are exemplified by the Windows of Hope (WOH) project (https://wohproject.com/), which entails a highly translational community genomics research programme founded by research supervisors Prof Andrew Crosby and Dr Emma Baple. The WHO study has enabled the identification of 25 novel genetic conditions to date, and the definition of >150 previously identified genetic disorders, which were not previously recognised in the community. This work has enabled immensely productive translational benefits for patients, their families, and the Amish community, and provided similar benefits for families with the same genetic conditions globally. All of the genetic discoveries we have made as part of our Amish genomic medicine programme have led to new genetic testing approaches, which have then been integrated into regional (serving Amish communities) and international diagnostic laboratories. Our research team have developed and delivered educational programmes on the nature, spectrum and causes of inherited disorders affecting the Amish, aimed at the community and also for healthcare service professionals. We have also developed online educational resources including a Massive Open Online Course (MOOC) entitled 'Genomic Medicine and Research: A Community Approach' (https://www.futurelearn.com/courses/community-genetics) and printed disease specific brochures to share research findings. Together, this work has increased diagnostic rates for genetic disease from <15% (2013) to >70% by 2022, reduced hospitalisations, prevented major neurological and physical impairments, and enabled estimated savings of >$100 million in community healthcare costs. The same aims and principles of our work with the Amish underpin all of our group's research studies in other communities and populations worldwide. In addition to the studies described in this thesis, I have also contributed to other clinical and genetic studies ongoing more widely as part of these research programmes at different stages of publication (see Appendix 6).

A common characteristic of genetic disorders that are present in Amish and Pakistani communities is their derivation from individual gene variants, introduced into the community by ancestral founder individuals (Figure 1.6). The subsequent transmission of these variants through the generations has resulted in a number of them accumulating to a substantial allele frequency and being associated with a notable burden of disease. This, combined with large nuclear family sizes and wider community genomic and environmental homogeneity greatly empowers genetic studies in these communities (see Section 1.6) and has greatly facilitated the identification of inherited neurodevelopmental disorders in these groups. Almost all autosomal recessive genetic conditions originally identified in the Amish or Pakistani communities have subsequently been reported in other communities worldwide. One important example is the discovery of *KPTN*-related disorder in nine individuals from the Amish community discovered by our research group in 2014 [333]. Chapter 4 details the subsequent identification of 27 further affected individuals with *KPTN*-related disorder from 15 distinct geographical regions internationally, highlighting the importance of studying inherited conditions within genetic isolates and the impact and relevance of these studies to other populations globally. The studies presented in Chapter 4 confirm that the *KPTN* variants p.(Ser259*) and p.(Met241_Gln246dup) associated with *KPTN*-related disorder are founder variants enriched within the Amish population, but also show that the p.(Met241_Gln246dup) variant has an European founder variant origin which has now been identified globally in 20 individuals with the condition.

The comprehensive genetic, clinical and molecular studies described in this PhD thesis deliver new specialist knowledge of rare autosomal recessive disorders of brain growth and development present in Amish and Pakistani communities. This has provided invaluable information for families with these disorders, particularly entailing diagnostic and clinical management benefits. For many families, confirmation of a specific genetic diagnosis may end the 'diagnostic odyssey' they often experience spanning many years, where diagnostic uncertainty may lead to extensive investigations in an attempt to identify a diagnosis for an affected individual's condition. A specific genetic diagnosis may commonly negate the requirement for exploratory clinical investigations that may be invasive, painful and expensive. Additionally, reaching a specific diagnosis enables accurate genetic counselling provision for

families who may wish to have further children. A good example of this entails defining a *CEP55* p.(Ile172Asnfs*17) variant as causative of a lethal fetal disorder in two siblings (Section 5.1), allowing accurate genetic counselling for the parents to be fully informed about the risk of recurrence in future pregnancies.

The identification of the specific causative genetic variants responsible for disease in a community context allows cost-effective and rapid carrier testing as well as early diagnostic testing strategies to be developed. Pathogenic variants identified by our team (and others) have been incorporated into a diagnostic testing panel for the Amish community that has been developed by the Exeter team in collaboration with industrial (PlexSeq, Cleveland) and CLIA (Clinical Laboratory Improvement Amendments)-certified (Wisconsin Clinical Laboratory) partners. This panel includes >150 founder genetic variants identified across the Amish communities and is available through the Wisconsin Clinical Laboratory to provide rapid and cost-effective testing for these conditions and is offered at cost to the Amish community. New disease-causing genetic variants identified within the Amish community can be incorporated within this panel, which has potential uses for preconception/prenatal carrier screening or newborn screening.

Enhanced neonatal and carrier screening has previously been offered to Anabaptist communities for specific founder variants associated with relatively common neonatal diseases in this community, including maple syrup urine disease (MSUD, OMIM #248600 (*BCKDHA* c.1312T>A)) and propionic acidaemia (PA, OMIM #606054 (*PCCB* c.1606A>G)) and phenylketonuria (PKU, OMIM #261600 (PAH c.284_286delTCA, c.782G>A, c.1066-11G>A)) [471-473]. Early detection is important to improve long-term outcomes for these conditions with dietary modifications and monitoring for complications [471, 474], although standard metabolic assays are known to miss a proportion of cases due to inconclusive findings in some individuals. A similar approach allowing early diagnosis of other neurodevelopmental disorders would enable early intervention and clinical management strategies to improve clinical outcomes of these conditions. Thus, the availability of this testing panel specifically for Amish founder variants has immense clinical utility for diagnostic testing, neonatal and carrier screening, as well as substantial cost saving outcomes for the wider community stemming from negating unnecessary diagnostic testing and targeted management.

Comprehensively delineating the clinical phenotype and progression of genetic neurodevelopmental disorders provides benefits for patients worldwide who may yet remain undiagnosed, with conditions caused by variants in these genes. Through these studies these genes will be incorporated in genetic testing panels (for example in PanelApp; (https://panelapp.genomicsengland.co.uk/) [123]) so individuals affected with these conditions worldwide are more likely to receive a diagnosis. This knowledge also allows clinicians to identify the most beneficial clinical management strategies for affected individuals through screening for specific clinical features or potential complications of a disorder. An excellent example are the clinical management guidelines for *KPTN*-related disorder deriving from the findings from these studies (Table 4.6), to aid clinicians to optimise management of patients worldwide. These guidelines are targeted at preventing potentially life-threatening complications, including monitoring infants for hypoglycaemia and optimising seizure management, as well supportive measures to optimise psychomotor development and level of functioning. Knowledge of effective supportive and symptomatic treatments allows their provision by local healthcare providers to improve outcomes, for example previous successful results from the use of specific antiepileptic medication in affected individuals with *KPTN*-related disorder to guide management of seizures in other affected individuals. A key finding in our studies of *KPTN*-related disorder is confirmation that this condition is an mTORopathy (Figure 4.11); as such mTOR inhibitors which have been used successfully in other mTORopathies (TSC, PMSE and Cowden syndromes [369, 371, 372] (see Section 4.2, Table 4.1)) may have genuine utility as a treatment for *KPTN*-related disorder. This is particularly exciting given the potential postnatal therapeutic window for this condition (Section 4.4, Figures 4.6 and 4.7) in which early treatment intervention may potentially prevent certain developmental aspects of the disorder. Clinical trials of mTOR inhibitors in individuals with *KPTN*-related disorder are in the planning stages by our group and collaborators which, if successful, would provide a new treatment approach for this condition worldwide. In collaboration with Professor Peter Crino's research group (University of Maryland) our group has been awarded long-term (7 year) NIH funding to pursue this research, and I am really pleased to be a part of this ongoing work with the group.

Further translational benefits of our work stem from regularly feeding back findings of our research through support group meetings to families with these conditions, as well

as educational meetings to local healthcare professionals targeted to improving knowledge of Amish inherited conditions aiming to develop effective management strategies and improve outcomes. To benefit communities globally, our research group has also cofounded family support groups, for example the 'KPTN Alliance' (https://kptnalliance.org/) which aims to support families with *KPTN*-related disorder by providing information and community connections internationally, working with medical professionals to share knowledge about the condition, support effective management and treatment and work with researchers globally to facilitate ongoing research to ultimately develop potential new treatments (for example mTOR inhibitors).

As detailed in Section 1.6 many genetic variants which are ultra-rare globally may become enriched in a community setting, allowing the clinical relevance of that particular variant to be interpreted. This includes benign gene variants which, due to limitations inherent in previous studies, have been incorrectly described in the literature as pathogenic. An excellent example of this is the sodium channel voltage-gated type IX (*SCN9A*) gene [424], in which variants were previously described as a cause of monogenic seizure phenotypes [475]. Stemming from the association of a rare *SCN9A* gene variant with the disease in a large Utah family comprising 21 individuals affected by febrile/afebrile seizures, the gene was considered a cause of this condition and included in genetic testing panels worldwide. However, through the investigation of the same *SCN9A* variant which has serendipitously accumulated in frequency in the Amish community, our research group was able to determine that this and other *SCN9A* gene variants are not a cause of seizure phenotypes. Equally importantly, the accumulation of founder gene variants in communities also enables the identification and clinical delineation of new genetic disorders which due to their ultra-rare nature globally, are enormously difficult to identify. Currently as only 6104 monogenic disorders have been described [68], a great deal of work still needs to be undertaken to identify the many ultra-rare monogenic diseases which remain yet to be discovered, as well as to define more complex genetic disease mechanisms (for example intronic, promotor enhancer/silencer element mutations, polygenic disease). Future studies utilising WGS will facilitate the detection of these variants as well as better detection of structural variants [64], to identify potential new causes of ultrarare neurodevelopmental disorders, although it is worth bearing in mind the pitfalls of these

technologies [63, 476, 477]. Studying genetic disorders in a community setting will enable important scientific advancements to be made in these areas, which would otherwise be difficult to achieve.

Defining new monogenic causes of neurodevelopmental disease in turn identifies molecular signalling pathways of crucial importance to brain development and function. Our studies of *TRAPPC10*-related disorder described in Chapter 3 provide an excellent example of this, by defining *TRAPPC10*-related disease in relation to other TRAPPopathy disorders arising through disruption of TRAPP II complex function, including effects on membrane trafficking via interaction with the TRAPPC2L subunit, as well as a previously unrecognised crucial interaction with TRAPPC9. *TRAPPC10*-related disorder shows significant clinical overlap with other TRAPPopathies, including features of microcephaly, severe global developmental delay/intellectual disability and pervasive behavioural abnormalities. This led to further studies identifying families with other potential TRAPPopathy disorders (TRAPPC1, TRAPPC8 and TRAPPC13 see Appendix 3) which also clinically fit with this family of conditions. Together these findings confirm the importance of the TRAPP II/III complexes for normal brain development and growth and define potential new avenues for functional evaluation of candidate VUS (see section 3.6).

A vastly expanding number of genes have been associated with autosomal recessive neurodevelopmental disorders in humans, for example around 400 genes have been associated with autosomal recessive microcephalic disorders alone (Figure 1.3). This highlights the complexity of the processes involved in brain development. Future challenges include using this scientific knowledge of the processes involved in brain development to develop novel therapies that are targeted at specific molecular pathways to prevent or counteract the disruption of these developmental processes to treat these neurodevelopmental disorders.

1 APPENDIX: Primer pairs, PCR conditions and plasmid structure

Primer	Primer sequences	Annealing temp °C	Product size (bp)
TRAPPC10 p.(Gly1131Valfs*19)	F: CCGCAGTGTTTTGTCAGACA R: TTGGGAGGAGTGTGCAGAAT	59	300
TRAPPC10 HAP1 knockout verification	F: GTCTGTAGTGGCCCTGGTTA R: TTAGACTAAGCCTGGCAGCC	57	835
TRAPPC10 HEK293 knockout verification	F: AGCGTAGTTATGATTTGGGGT R: GCCAAGGAATGAAGGGACAA	59	540
KPTN p.(Met241_Gln246dup) p.(Ser259*)	F: GCGGAAAGTCGATGGTTTCA R: CCAAGAATGCAGAGTTGCCTC	56	223
CEP55 p.(Ile172Asnfs*17)	F: GGCTCCAAACTGCTTCAACTCATC R: GTCACAGCACTAATGTTGCCAATTTC	63	154
INPP4A p.(Leu46*)	F: GAGGAGCAGGGGTTTGGTTG R: GGGCTCCGTCTTCTCAGAAA	60	388

Table A1.1: Primer pairs and PCR conditions

Primer	Primer sequences	Annealing temp °C	Product size (bp)
TRAPPC10	F: CTGACTGAATCTGATGAGCA R: TTGTCTACTGACAGGCTGTC	52	278 without introns

Table A1.2: cDNA primer pairs
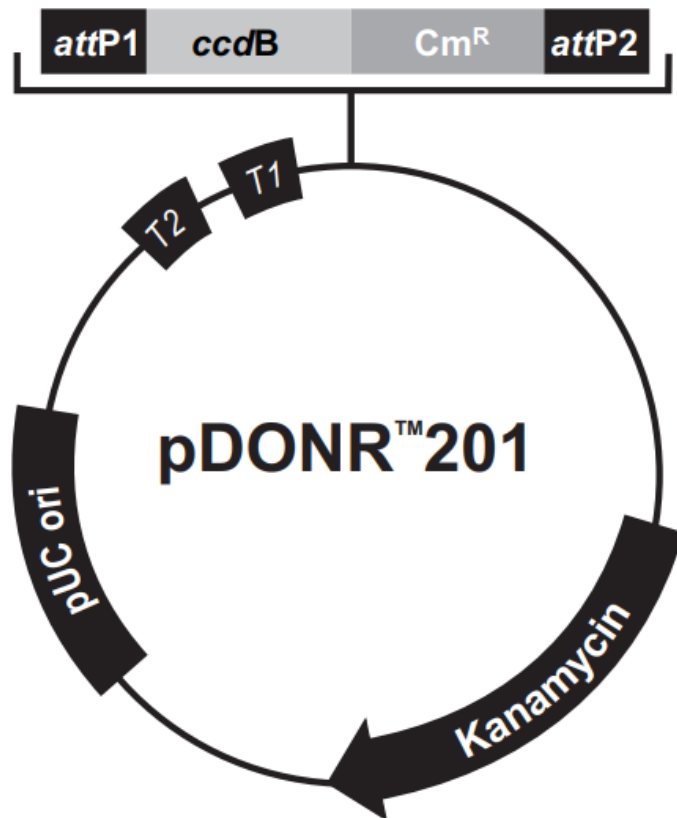
Figure A1.1: Elements of pDONR™201

Reproduced from Gateway® pDONR™Vectors User Manual 2007.
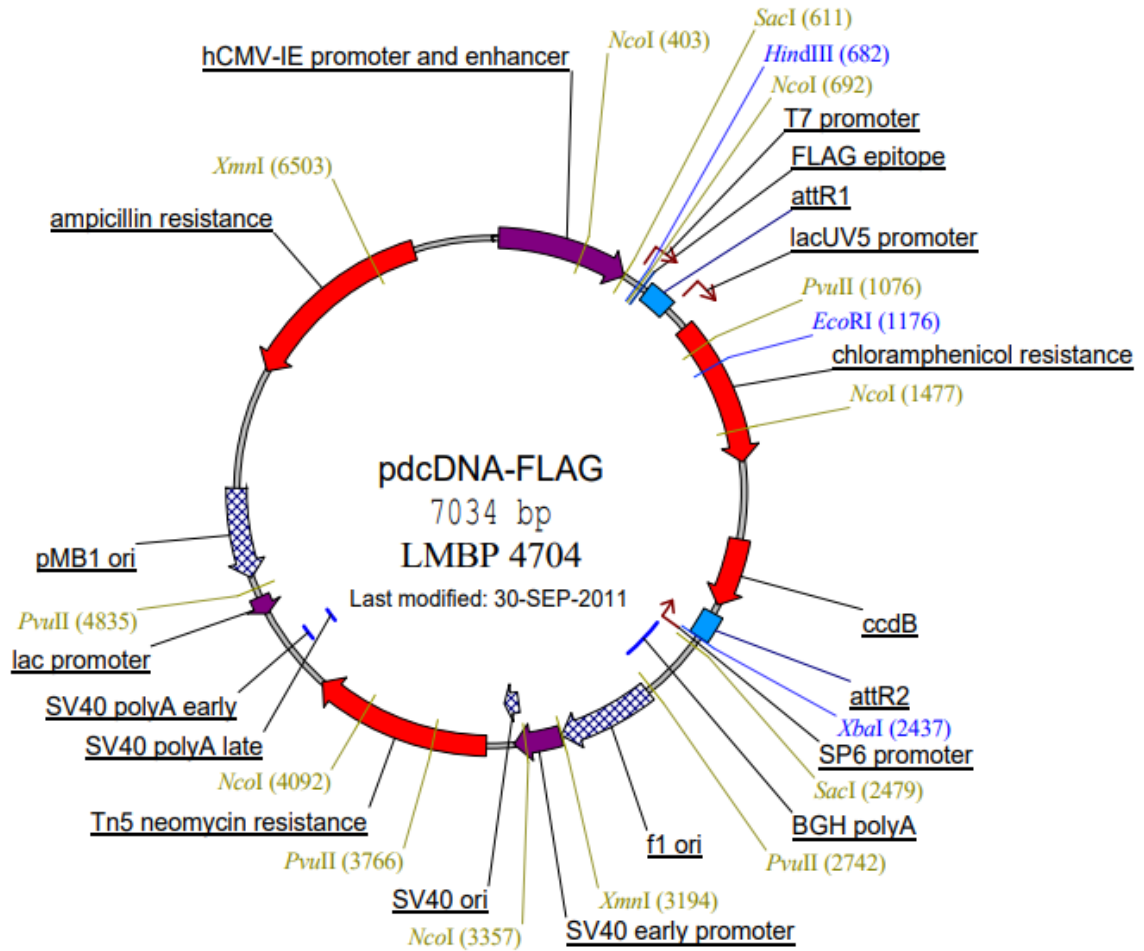
Figure A1.2: Structure and restriction sites of FLAG plasmid

Reproduced from the Belgian Coordinated Collections of Microorganisms (https://bccm.belspo.be/) 2022.

2 APPENDIX: Methods and results of molecular and mouse studies of *TRAPPC10*-related neurodevelopmental disorder

A2.1 Methods

A2.1.1 Functional studies

The following studies were performed by Hashem Almousa, Ana Maria Hincapie, Dr Miroslav Milev and Dr Djennan Saint-Dic.

Construction of TRAPPC10$^{-/-}$ cell lines

Guide RNA (gRNA) design and cloning (CRISPR plasmid pSPCas9(BB)-2A-Puro (PX459) V2.0, Addgene #62988) was performed as described previously [478], with primers: 5'-CACCGCATCTTCGGAGCCCGGCCAT-3' and 5'-AAACATGGCCGGGCTCCGAAGATGC-3'. HEK293 and HAP1 cells were transfected (70% confluency) using JetPRIME with 500ng DNA expressing the gRNA. Following selection (puromycin) and colony expansion, genomic DNA was extracted. The *TRAPPC10* gene region of interest was amplified using oligonucleotide pair: 5'-AGCGTAGTTATGATTTGGGGT-3 and 5'-GCCAAGGAATGAAGGGACAA-3. The resulting PCR product (~540bp, agarose gel electrophoresis) was sequenced. Knockout verification was performed by western analysis.

Membrane trafficking assay

Cells were infected with virus expressing VSVG-GFP-tsO45 (one hour, 37°C). In rescue experiments, cells were transfected with either FLAG-TRAPPC10, FLAG-TRAPPC10 p.Gly1131Valfs*19 or FLAG-TRAPPC10 p.Pro929Leu constructs, then infected with VSVG–GFP ts045 (one hour, 37°C). Cells were maintained at 40°C overnight. Cycloheximide was added (final concentration, 10μg/ml), prior to shifting cells to 32°C. Time-lapse microscopy started three minutes after temperature shift to allow time to select fluorescent cells (Nikon inverted confocal microscope) as previously described [310]. Movies used for quantitative fluorescence analysis were not subjected to processing. Integrated fluorescence intensity at the Golgi region (defined by the region of perinuclear intensity seen 20-40 minutes after temperature shift) and from whole cell was measured using ImageJ (measurements obtained every 5 minutes). The ratio between fluorescent intensities within the Golgi region and whole

cell was generated for each time point. The kinetics of VSVG-GFP-tsO45 trafficking represent change in that ratio over time (0 to 90 min).

Cell lysis and size exclusion chromatography

Cells were grown in either DMEM (for HEK293) or RPMI 1640 (for lymphoblastoid cells) medium containing 10% fetal bovine serum (FBS). Cells were lysed in a solution containing 50mM Tris pH 7.2, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 1% Triton X-100 and protease inhibitor cocktail (EDTA-free; Roche). A total of 2-5 mg of protein was fractionated on a Superose 6 Increase 10/300 GL column at a flow rate of 0.4 ml/min. Fractions of 0.5 ml were collected in wash buffer (50 mM Tris pH 7.2, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 0.1% Triton X-100) and probed with the indicated antibodies.

Yeast two hybrid assay

Open reading frames (ORFs) encoding TRAPP proteins were cloned (pGADT7 and pGBKT7 vectors; Clontech). Plasmids were transformed into yeast cells (AH109, Y187). Diploids containing the respective TRAPP ORFs were produced by mating and selecting on synthetic complete medium (SC) lacking leucine and tryptophan. Interactions were assessed on SC medium lacking leucine, tryptophan, histidine and adenine. Serial cell dilutions were spotted onto solid medium. Quantification of interaction was assessed using a β-galactosidase assay employing ONPG and normalized to total protein.

Antibodies

Commercially available antibodies used in this study were: anti-TRAPPC2 (rabbit polyclonal, homemade), anti-TRAPPC2L (mouse monoclonal, Santa Cruz sc-377322), anti-TRAPPC3 (rabbit polyclonal, homemade), anti-TRAPPC8 (rabbit polyclonal, Abcam ab122692), anti-TRAPPC9 (rabbit polyclonal, LS Bio LS-C750497), anti-TRAPPC10 (mouse monoclonal, Santa Cruz sc-101259), anti-TRAPPC12 (rabbit polyclonal, homemade), anti-FLAG (mouse monoclonal, Sigma F1804), anti-α-Tubulin (mouse monoclonal, Sigma T6199).

A2.1.2 Mouse studies

Studies were performed by Dr Stephan C Collins, Dr Christine Rowley, Dr Valerie E Vancollie and Dr Christopher J Lelliott.

Trappc10 knockout mouse model

The mouse model was generated by homologous recombination in embryonic stem cells using the Knockout-first allele method [479], adopting a strategy identifying an exon common to all transcripts (exon 14), upstream of which a LacZ cassette was inserted (Figure A2.1). Exon 14 of the *Trappc10* allele, flanked by *loxP* sequences bilaterally, was deleted using a Cre recombinase that recognizes *loxP* sites, producing the *Trappc10*[tmb(EUCOMM)Wtsi] knockout allele. Mice were phenotyped by the Mouse Genetics Project (MGP) pipeline at the Wellcome Sanger Institute, UK. The care and use of mice was carried out in accordance with UK Home Office regulations (license number 80/2076), UK Animals (Scientific Procedures) Act of 1986 under UK Home Office license (80/2076) that approved this work, which was reviewed regularly by the Wellcome Sanger Institute Animal Welfare and Ethical Review Body. Animal care was as previously described by White *et al* (2013) [380]. Mice were fed a standard chow diet using the autoclavable mouse breeder diet 5021 (www.labdiet.com). After weaning, animals were housed three to four mice per cage with WT controls housed separately, in a specific pathogen-free environment in individually ventilated cages under 12/12 light/dark cycle with temperature-controlled conditions and free access to food and water with hardwood bedding. All animals were regularly monitored for health and welfare concerns and were additionally checked prior to and after procedures. Mice of both genders were weighed between 4 and 16 weeks of age.

Figure A2.1: Allelic construction showing targeting of critical exon 14 in the Trappc10$^{tm1b(EUCOMM)Wtsi}$ mouse

Data courtesy of Dr Binnaz Yelcin.

Mouse neuroanatomical studies

Neuroanatomical studies were carried out using three homozygous *Trappc10*[-/-] and 498 baseline wild type (WT) mice on a C57BL/6N pure genetic background at 16-weeks of age as previously described [324, 480]. We used a recognized statistical model (Gpower) validated for comparison of modest numbers (n=3) of animals to evaluate neuroanatomical defects with an effect size of 10% or more with 80% detection power [480]. Mice were anaesthetized with Ketamine (100 mg/kg, intraperitoneally) and Xylazine (10 mg/kg, intraperitoneal), blood collected via the retro-orbital route and death confirmed before the brains were dissected and fixed in 4% buffered formalin for 48 hours, then transferred to 70% ethanol. Samples were embedded in paraffin using an automated embedding machine (Sakura Tissue-Tek VIP) and cut at 5µm thickness (sliding microtome, Leica RM 2145) to obtain coronal brain region at Bregma +0.98 mm and Bregma -1.34 mm according to the Allen Mouse Brain Atlas [481]. Sections were stained with 0.1% Luxol Fast Blue (Solvent Blue 38; Sigma-Aldrich) and 0.1% Cresyl violet acetate (Sigma-Aldrich) and scanned (Nanozoomer 2.0HT, C9600 series) at 20× resolution. Sixty-three brain parameters made up of area and length measurements as well as cell level features, were taken across the two coronal sections. Co-variates, for example sample processing dates and usernames, were collected at every step of the procedure and used to identify data drifts. Using in-house ImageJ plugins, an image analysis pipeline was used to standardize measurements of areas and lengths. Images were quality controlled for the accuracy of sectioning relative to the reference atlas and controlled for asymmetries and histological artefacts. At Bregma +0.98mm brain structures assessed were: 1) the total brain area, 2) the lateral ventricles, 3) the cingulate cortex, 4) the genu of the corpus callosum, 5) the caudate putamen, 6) the anterior commissure, 7) the piriform cortex, 8) the primary motor cortex and 9) the secondary somatosensory cortex. At Bregma -1.34mm, a maximum of 14 brain structures were assessed: 1) the total brain area, 2) the lateral and third ventricles, 3) the retrosplenial granular cortex, 4) the corpus callosum, 5) the amygdala, 6) the piriform cortex, 7) the internal capsule, 8) the optic tract, 9) the mammillothalamic tract, 10) the fimbria of the hippocampus, 11) the habenular nucleus, 12) the hippocampus, 13) the primary motor cortex and 14) the secondary somatosensory cortex. All samples were also systematically assessed for cellular ectopia (misplaced neurons). Data were analysed

using a linear mixed model framework to determine whether a brain region is associated with neuroanatomical defect or not.

For the neuroanatomical assessment of *Trappc9⁻/⁻* mice, the brains were cut into halves along the sagittal midline, and embedded separately in paraffin. The section of interest was determined as corresponding to Figure 106 of the Mouse Brain Atlas (Lateral +0.60 mm)[482]. Brains were sectioned to match the defined section at a thickness of 5µm. The staining, scanning, and quality control steps were identical to the coronal procedure used for the neuroanatomical assessment of *Trappc10⁻/⁻* mice. A total of 40 brain morphological parameters including 25 areas, 14 lengths and one number, were measured for each brain. These parameters encompassed 22 distinct brain regions.

Mouse body composition and clinical blood chemistry tests

Non-fasted mice at 16 weeks of age were terminally anaesthetised with ketamine/xylazine and blood was collected into lithium/heparin coated tubes via the retro-orbital sinus. Plasma was analysed for the following 27 parameters on an Olympus AU400: 1) for electrolytes: sodium, potassium and chloride; 2) for non-fasted metabolic panel: glucose, fructosamine, triglycerides, cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), non-esterified free fatty acids (NEFAC) and glycerol; 3) for thyroid: total thyroxine and adiponectin; 4) pancreatic enzyme: amylase; 5) liver/muscle: alanine aminotransferase, alkaline phosphatase, creatine kinase, aspartate aminotransferase and total bilirubin; 6) protein parameters: total protein and albumin; 7) kidney: creatinine and urea; and 8) minerals and iron: calcium, magnesium, iron and phosphate. For body composition, 14 week mice were imaged on a dual energy X-ray absorptiometry machine (Lunar PIXImus II). This generated an image of the entire mouse and provided bone mineral, body composition and morphometric data.

A2.2 Results
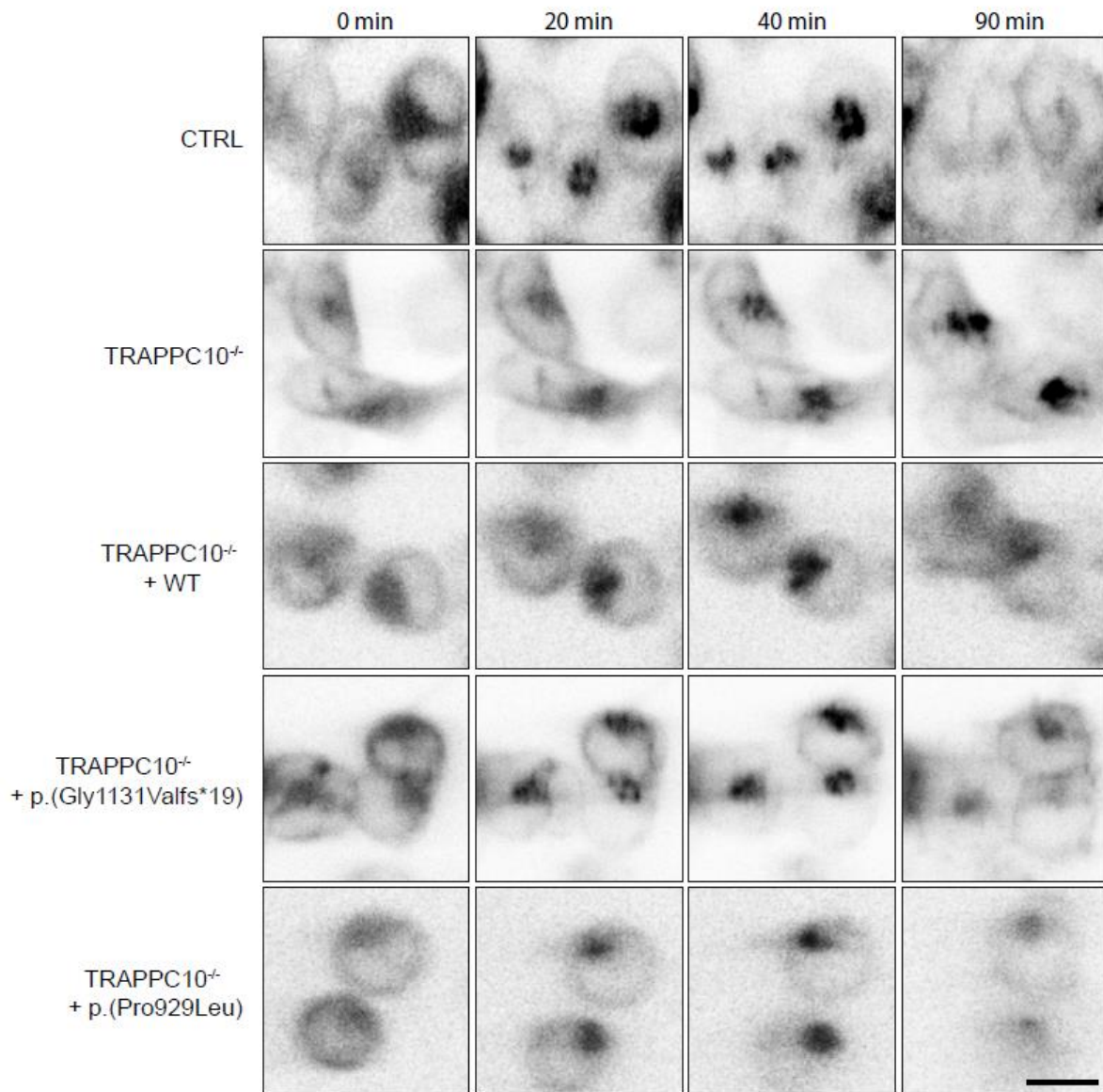
Figure A2.2: Representative images for VSV-G trafficking

Wild type HEK293 cells or *TRAPPC10*[-/-] cells either not transfected or transfected with FLAG-tagged wild type TRAPPC10 or one of the TRAPPC10 variants indicated were infected with VSVG-GFP ts045 4 hours after transfection. After an overnight incubation at 40°C, the cells were shifted to 32°C and imaged every minute. Representative images at 0, 20, 40 and 90 minutes are shown. Data courtesy of Dr Michael Sacher.
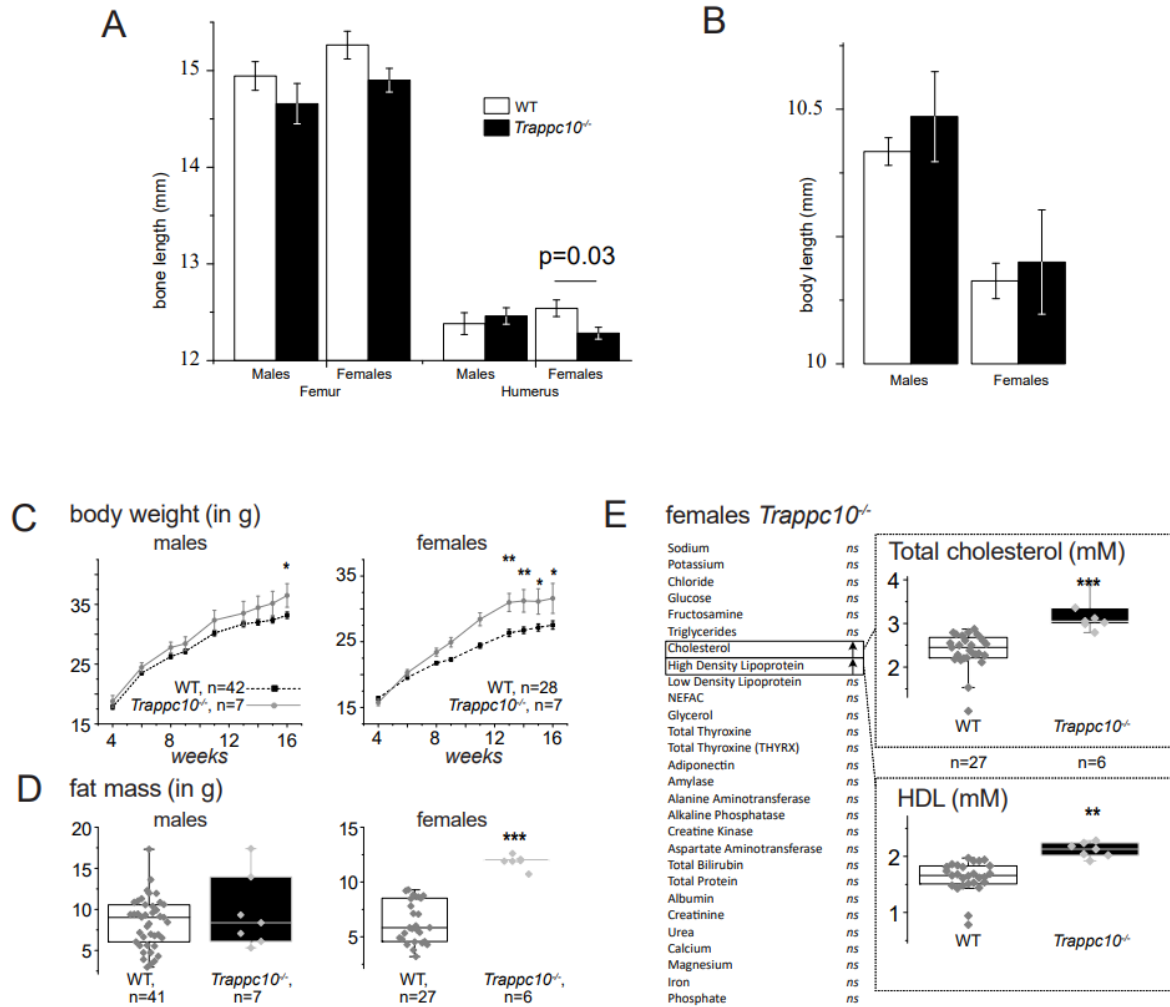
Figure A2.3: Whole-body phenotype analysis of *Trappc10* knockout mouse

(**A**) Bone length for long bones (Femur and Humerus) in both sexes (n=7 for all groups) (**B**) Body length for male *Trappc10⁻ᐟ⁻* (n=7) and WT (n=41) and females *Trappc10⁻ᐟ⁻* (n=7) and WT (n=27). (**C**) Body weight curves in grams of male and female *Trappc10⁻ᐟ⁻* mice, between 4 and 16 weeks of age. (**D**) Fat body composition in grams of male and female *Trappc10⁻ᐟ⁻* mice at 16 weeks of age. (**E**) *Left:* List of 27 assessed clinical blood chemistry parameters and association in female *Trappc10⁻ᐟ⁻* mice at 16 weeks of age. *Right:* Box plots with raw data points showing results for levels of total cholesterol and high-density lipoprotein (HDL) in millimoles per litre in female *Trappc10⁻ᐟ⁻* mice at 16 weeks of age. Statistical analyses were performed with GraphPad Prism 8.0.2, using two-tailed Student's t-tests of equal variances. *p<0.05 **p<0.01 ***p<0.001. Arrows indicate directionality of effect and "ns" indicates not significant (p-value>0.05). Data courtesy of Dr Binnaz Yelcin.
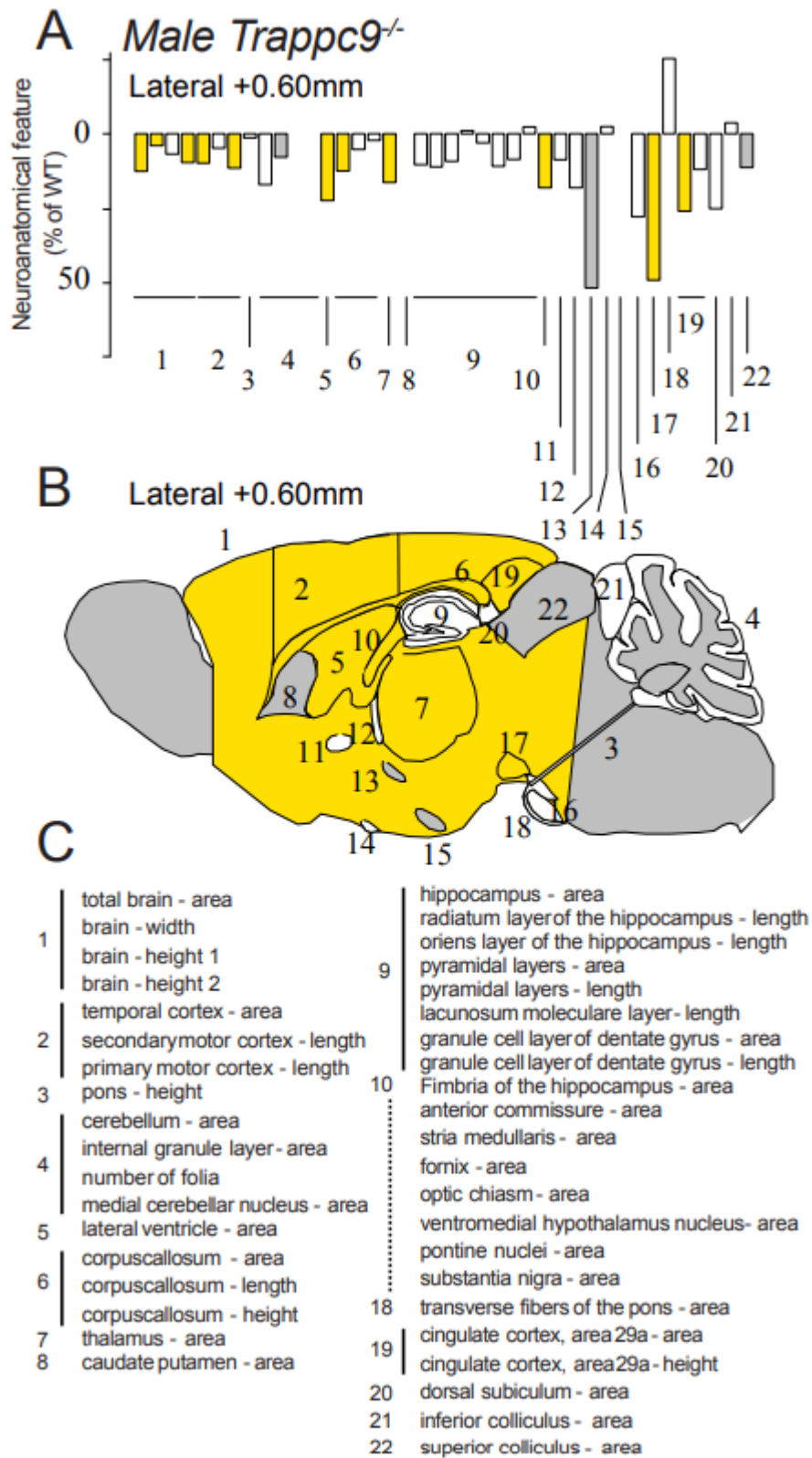
Figure A2.4: Neuroanatomical assessment of *Trappc9* knockout mouse

(**A**) Histograms of percentage change relative to *Trappc9*[+/+] (set as 0) for each of the measured parameters. (**B**) Schematic representation of the 22 brain regions quantified at lateral +0.60 mm on

sagittal section from *Trappc9$^{+/+}$* (n=4) and *Trappc9$^{-/-}$* (n=4) mice. Coloured regions indicate the presence of at least one significant parameter within the brain region at the 0.05 level. White indicates a p-value > 0.05, grey shows not enough data to calculate a p-value. (**C**) List of assessed brain parameters. Data courtesy of Dr Binnaz Yelcin.

3 APPENDIX: Identifying candidate new and consolidating previously unconfirmed TRAPPopathy disorders

A3.1 Introduction

Sixteen human TRAPP complex proteins have been identified to date (Figure 3.1). Of these, sequence variants within 11 have been associated with human disease (Table 3.1). This section describes candidate novel TRAPPopathy disorders pertaining to three of the remaining five TRAPP complex subunits (*TRAPPC1-*, *TRAPPC5- and TRAPPC8*-associated neurodevelopmental conditions) as well as providing new genetic evidence to consolidate previous findings relating to two other TRAPP complex proteins (*TRAPPC3* and *TRAPPC14*) in which candidate single families with human disease due to variants in these genes have previously been identified (Table 3.1).

A3.2 Methods

A3.2.1 Literature review

A detailed literature review was conducted to identify which TRAPP complex proteins have been associated with clinical disorders (Table 3.1). Sources used for literature review included Pubmed (https://pubmed.ncbi.nlm.nih.gov/), Google Scholar (https://scholar.google.com/), OMIM (https://www.omim.org/) and Human Gene Mutation Database (HGMD) Professional (https://digitalinsights.qiagen.com/product-login/). Search terms included TRAPPC gene names and previous gene symbols/names as listed by HUGO Gene Nomenclature Committee (HGNC) (https://www.genenames.org/), as well as the names of equivalent subunits in yeast.

A3.2.2 Data analysis

An application for a DDD CAP (#303) was approved and data transferred from the Wellcome Sanger Institute DDD team. Genetic data was provided for all known TRAPP complex genes and reviewed for *TRAPPC1*, *TRAPPC3*, *TRAPPC3L*, *TRAPPC5*, *TRAPPC6A*, *TRAPPC8*, *TRAPPC10*, *TRAPPC13* and *TRAPPC14*. Variants were filtered according to autosomal recessive inheritance (homozygous and compound heterozygous) with variant prioritisation criteria as described in Table 2.7.

Individuals with potentially pathogenic biallelic variants were identified, so that associated clinical phenotypic data could be reviewed and recorded. For biallelic missense variants individuals were excluded if the clinical phenotype was not compatible with a neurodevelopmental disorder similar to other known TRAPPopathies. Exome data was reviewed to identify any other possible causative variants located genome-wide in these individuals.

An application for access to the Genomics England (GE) research environment (https://re.extge.co.uk/ovd//) was granted to review data generated from the 100,000 genomes research project. Gene specific data was produced using a script written for the GE research environment by Dr James Fasham (University of Exeter). Data was first filtered for phenotypes compatible with a neurodevelopmental disorder and then variants analysed and filtered as above. Any protein truncating variants were further assessed for zygosity and clinical phenotype. Genome data was reviewed to identify any other possible causative variants (Tier 1 or 2) in these individuals.

A3.2.3 Collaboration to identify other families

Further potential TRAPPopathy families were identified through GeneMatcher [244]. Pathogenic variants within TRAPP complex proteins were reviewed in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) to identify any further patients with disorders caused by variants in TRAPP complex genes, and collaborating groups contacted for further details.

Detailed genetic and clinical information was gathered from collaborating groups as described in section 2.2.2. Where patients had given specific informed consent, clinical photographs and videos were shared and reviewed (section 2.2.1).

For TRAPP complex variants published with limited genetic and clinical data in large 'candidate disease gene finder' papers, the authors were contacted for further genetic and clinical information about these families.

A3.3 Results

A3.3.1 Literature review

Clinical phenotypes in affected individuals with variants in TRAPP complex proteins associated with known TRAPPopathies were identified following literature review and the results are summarised in Table 3.1. Table A3.1 provides detailed phenotypic information of previously reported TRAPPopathy disorders alongside previously undescribed TRAPP complex protein disorders, allowing comparison of clinical features. *TRAPPC2*, *TRAPPC2L*, *TRAPPC4*, *TRAPPC6B*, *TRAPPC9*, *TRAPPC11* and *TRAPPC12* genes (white columns) have a well-established association with a specific disorder identified in multiple families, and have an associated OMIM number. *TRAPPC3*, *TRAPPC6A*, *TRAPPC10* and *TRAPPC14* genes (green columns) have only a single family reported in the literature with suspected pathogenic variants associated with a neurodevelopmental disorder.

TRAPP subunit	TRAPPC1	TRAPPC2	TRAPPC2L	TRAPPC3	TRAPPC4	TRAPPC5	TRAPPC6A
OMIM #	-	313400	618331	-	618741	-	-
Disorder	-	Spondylo-epiphyseal dysplasia tarda (SEDT)	Encephalopathy, progressive, early-onset, with episodic rhabdomyolysis	-	Neurodevelopmental disorder with epilepsy, spasticity, and brain atrophy	-	-
Inheritance	AR	XLR	AR	AR	AR	AR	AR
Age of onset	Infancy	5-10 years	Infancy	NK	Neonatal	NK	
Neurological Features							
Intellectual disability	+	-	+	+	+ (severe)	+	+
Global developmental delay	+	-	+ (severe, regression)	+	+ (severe)	+ (severe)	+
Speech delay	+	-	+ (non-verbal)	+	+ (non-verbal)	+	+
Microcephaly	+	-	+ post-natal	+	+ progressive	+	-
Childhood hypotonia	+	-	NK	NK	+	NK	NK
Seizures	+ (1 individual)	-	+ (encephalopathy)	-	+ (early-onset)	+	-
Tetraplegia	-	-	+	-	+	NK	-
Spasticity	+ (1 individual)	-	-	NK	+	NK	-
Dystonia/choreoathetosis	-	-	-	NK	+	+	-
Myopathy	+ (1 individual)	-	+ (muscle biopsy)	NK	-	NK	+ (single patient)
Tremor	NK	-	-	NK	+	NK	-
Gait abnormalities	Non-ambulatory, ataxia	-	Non-ambulatory	NK	Non-ambulatory	NK	NK
Behavioural abnormalities	Autism, self-aggression	-	-	Social communication problems	-	NK	NK
Corpus callosum hypoplasia	+	-	-	-	+	NK	NK
Cerebral atrophy	+	-	+ (progressive)	-	+	NK	NK
Cerebellar atrophy	-	-	-	+	+	NK	NK
Other MRI features	Hypomyelination, olfactory	-	Delayed myelination	-	Ventriculomegaly, optic atrophy	NK	NK

	nerve/pons hypoplasia						
Skeletal features							
Short stature	NK	+	NK	+	NK	NK	-
Skeletal dysplasia	-	+	-	-	-	-	-
Raised creatine kinase	+	-	+ (episodic rhabdomyolysis)	NK	NK	NK	NK
Obesity	NK	-	-	+	-	NK	-
Polydactyly	-	-	-	+ (postaxial)	-	-	+ (postaxial)
Scoliosis/kyphosis	NK	+	NK	NK	+ (severe)	NK	-
Cardiac features	NK	-	-	Dextrocardia, complex cardiac disease	-	-	-
Audiological features	NK	-	-	NK	SNHL	NK	-
Ophthalmic features	NK	Corneal opacities	Cortical visual impairment	Nystagmus, RP	Cortical visual impairment, nystagmus, cataracts	Cerebral visual impairment, blindness	-
Other	Multiple verrucae Raised CK, gastrostomy	Joint pain, arthritis	Episodic exacerbation with illness	Penoscrotal hypospadias	Gastrostomy, severe infection, early death	Neurogenic bladder	Facial dysmorphism
References	-	[140, 277, 279]	[276, 280]	[281]	[282-286]	-	[287]

TRAPP subunit	TRAPPC6B	TRAPPC8	TRAPPC9	TRAPPC10	TRAPPC11	TRAPPC12	TRAPPC14
OMIM #	617862	-	613192	-	615356	617669	618351
Disorder	Neurodevelopmental disorder with microcephaly, epilepsy, and brain atrophy	-	Mental retardation, autosomal recessive 13	-	Muscular dystrophy, limb-girdle, autosomal recessive 18	Encephalopathy, progressive, early-onset, with brain atrophy and spasticity	?Microcephaly 25, primary, autosomal recessive
Inheritance	AR	AR	AR	AR	AR	AR	AR
Age of onset	Early infancy	NK	Infancy	Infancy	Early childhood	Birth	Birth
Neurological Features							
Intellectual disability	+ (severe)	+	+ (moderate-severe)	+ (severe)	+	+	+
Global developmental delay	+ (severe)	+ (regression)	+	+	+	+ (severe, regression)	+
Speech delay	+ (non-verbal)	+	+ (few words)	+ (some non-verbal)	+	+	+ (few words)
Microcephaly	+ (progressive)	+	+ (postnatal)	+	+	+	+
Hypotonia	+	NK	+	+	+	+ (truncal)	NK
Seizures	+ (generalised)	+ (febrile)	+ (rare)	+	+	+	-
Tetraplegia	-	+	-	-	-	+	-
Spasticity	-	+	-	-	-	+ (appendicular)	-
Dystonia/choreoathetosis	-	-	-	-	+	+	-
Myopathy	NK	NK	-	-	+ Gower's sign	-	NK
Tremor	+	NK	-	-	+	-	NK
Gait abnormalities	Ataxia	NK	-	Waddling gait	Ataxia, non-ambulatory	-	NK
Behavioural abnormalities	Stereotypies, poor social interaction	NK	Hyperactivity, happy disposition	Aggressive episodes, autism	-	-	ADHD
Corpus callosum hypoplasia	+	NK	+	+	+	+	+
Cerebral atrophy	+	NK	+	-	+	+	
Cerebellar atrophy	+	NK	+	-	+	-	-

Other MRI features	Brainstem atrophy, ventriculomegaly	Leukodystrophy	White matter abnormalities	-	Hypomyelination	Pons hypoplasia, simplified gyri, ventriculomegaly	Reduced white matter volume
Skeletal features							
Short stature	+	+	-	+	+	NK	+
Skeletal dysplasia	-	-	-	-	-	-	-
Raised creatine kinase	NK	NK	NK	NK	+	NK	NK
Obesity	-	+	+	-	-	-	-
Polydactyly	-	-	-	-	-	-	-
Scoliosis/kyphosis	NK	+	+	-	+	+	-
Cardiac features	-	-	-	-	-	-	-
Audiological features	NK	Bilateral sensorineural loss	-	Otitis media	-	Hearing loss	-
Ophthalmic features	Nystagmus, strabismus	Myopia, retinal dystrophy, pectus carinatum	-	Strabismus	Cataracts, strabismus, myopia, alacrima	Cortical visual impairment, optic atrophy	-
Other	Dysphagia	Hepatomegaly	Facial dysmorphism	Multiple verrucae	Hepatomegaly, poor feeding	Dysphagia, poor feeding, neurogenic bladder	Tethered spinal cord
References	[288-291]	-	[292-306]	[307]	[272, 273, 308-316]	[317-319]	[320]

Table A3.1: Clinical features reported in affected individuals with variants in TRAPP complex proteins

All display autosomal recessive inheritance except *TRAPPC2*-associated spondylo-epiphyseal dysplasia tarda (SEDT) which displays X-linked recessive inheritance. White columns indicate well characterised TRAPPopathy genes with multiple publications confirming disease causation. Green columns indicate previously identified potential TRAPPopathy genes with limited evidence published for disease causation in single families. Orange columns indicate TRAPP complex genes with no previously published disease association. '+' indicates the presence and '-' indicates absence of a feature, NK; not known.

A3.3.2 Identification of candidate novel TRAPPopathy disorders

A total of seven families with affected individuals with neurodevelopmental disorders potentially caused by biallelic variants in TRAPP complex associated proteins not previously associated with a clinical disorder were identified, described below.

TRAPPC1

TRAPPC1 is a widely expressed core TRAPP complex protein of 145 amino acids [483], with high levels of expression in the frontal cortex and other brain regions (GTex, Broad Institute 2022). TRAPPC1 has been shown to interact with the TRAPPC3/TRAPPC6B heterodimer within the core TRAPP complex [484]. A single protein domain has been identified in TRAPPC1: a synbindin-like family domain (Figure A3.1). Synbindin/TRAPPC4 is a syndecan-2 ligand on the surface of dendritic spines that receive the vast majority of excitatory synapses [329].

Three families with rare, predicted deleterious biallelic variants in *TRAPPC1* discovered as the potential cause of a similar neurodevelopmental disorder were identified. Family 1 comprises a nine year old affected Spanish female with a neurodevelopmental phenotype including features of severe intellectual disability/GDD (non-verbal and non-ambulatory) with significant hypotonia, possible myopathy, visual impairment and autistic features. Investigations identified persistently raised creatine kinase (CK), with minor non-specific changes on muscle biopsy (Table A3.2). MRI neuroimaging identified delayed myelination, hypoplastic corpus callosum and optic nerves. WES identified compound heterozygous *TRAPPC1* NM_021210.5 variants (Chr17:g.7930680GCA>G, c.362_363del; p.(Val121Alafs*3) [hg38], gnomAD allele frequency 0.0001662 with no homozygotes listed) and (Chr17:g7931758_7931766del, c.64_72del; p.(22_24del), absent in gnomAD) that segregate within the family (Figure A3.1). These variants are classified as pathogenic and likely pathogenic respectively according to ACMG/ACGS classification criteria [401] (Table A3.3). The p.(Val121Alafs*3) variant is one of ten LOF *TRAPPC1* variants listed in gnomAD, none are homozygous.

FAMILY	FAMILY 1 II:1	FAMILY 2 II:1	FAMILY 3 II:1
Ethnicity	Spain	Germany	UK-Asian
Genotype	p.(Val121Alafs*3)/p.(22_24del)	p.(His98Pro)/p.(His72Arg)	c.170+5G>A/c.170+5G>A
Age at assessment (years)	9	4	11
Sex	F	F	F
GROWTH			
Birth gestation (weeks)	38	NK	NK
Birth weight kg (SD)	NK	NK	NK
Birth OFC cm (SD)	NK	Normal	NK
Height cm (SD)	NK	NK	NK
Weight kg (SD)	NK	NK	NK
OFC cm (SD)	NK	Microcephaly	Microcephaly
DEVELOPMENT			
Intellectual disability	+ (severe)	+	+ (severe)
Global developmental delay	+	+	+
Speech impairment	Non-verbal, uses basic signs	+	+
Motor milestones	Non-ambulatory, stands with support, bottom shuffles	Delay	Non-ambulatory, gross and fine motor delay
Vision	-	NK	NK
Hearing	-	NK	NK
NEUROLOGY			
Tone	Hypotonia (Increased centrally)	NK	NK
Muscle weakness	+ (possible myopathy)	NK	NK
Spasticity	+ with brisk reflexes	NK	NK
Cerebellar signs	-	NK	Ataxia
Seizures	-	+ onset 13 months	-
Neuroimaging	Thin corpus callosum, hypoplastic olfactory nerves, hypomyelination	Progressive brain atrophy, microcephaly	Hypoplasia of the pons
EEG	Slow background rhythms for age, no epileptiform activity	NK	NK

Behavioural abnormalities	Autism, self-injurious behaviour, body rocking, stereotypical movements	NK	NK
DYSMORPHIC FEATURES			
Facial features	Long palpebral fissures, hypertelorism, high arched eyebrows, bulbous nasal tip, large ears, high arched palate, widely spaced teeth	NK	Micrognathia
Limbs	Single palmar crease on right hand	NK	NK
OTHER	Fundoplication and gastrostomy insertion Chronically raised CK (600-2000)	Nil	Multiple verrucae on fingers and toes

Table A3.2: Clinical features of individuals with biallelic variants in *TRAPPC1*

'+' indicates the presence and '-' indicates absence of a feature. Abbreviations: SD; standard deviation, OFC; occipitofrontal circumference, CK; creatine kinase, EEG; electroencephalogram, NK; not known.
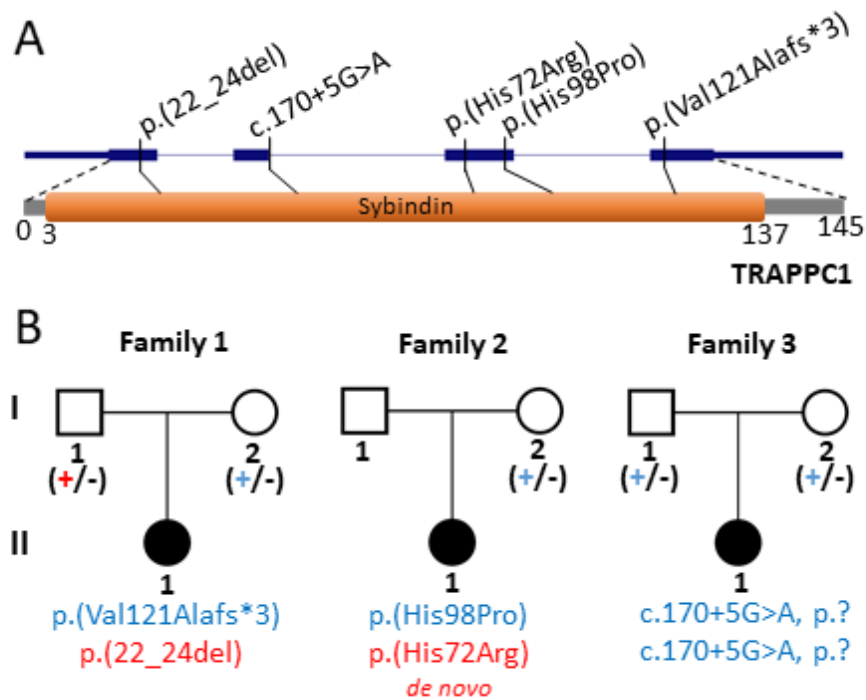
Figure A3.1: *TRAPPC1* variants and pedigrees of affected individuals

(**A**) Schematic of the *TRAPPC1* gene (blue) showing coding exons in wide blue bars, non-coding exonic regions as narrower blue bars and introns as a blue line. The TRAPPC1 protein product is shown below including known protein domain synbindin (orange) [341]. The location of *TRAPPC1* variants are indicated by black lines. (**B**) Pedigrees of three female singleton affected individuals with biallelic *TRAPPC1* variants segregating for an autosomal recessive disorder; '+' indicates the presence of a variant and '-' indicates absence.

Family 2 comprises a four-year-old female from Germany, affected by a neurodevelopmental disorder from age 3 months, characterised by global developmental delay, seizures, secondary microcephaly and progressive brain atrophy identified on MRI neuroimaging (Table A3.2; further clinical details pending). WES identified compound heterozygous *TRAPPC1* variants, maternally inherited (Chr17:g.7931027T>G, c.293A>C; p.(His98Pro) [hg38], gnomAD; absent) and a *de novo* variant (Chr17:g.7931105T>C, c.215A>G, p.(His72Arg)) *in trans* (Figure A3.1, Table A3.3); both variants are within the synbindin domain and are predicted to be deleterious by multiple missense prediction tools (REVEL scores 0.725 and 0.804 respectively).

Family 3 is of British-Asian origin with an 11 year old affected female with features of intellectual disability, GDD, microcephaly, severe ataxia with an inability to walk (Table A3.2) and pons hypoplasia on MRI neuroimaging. Trio WGS (100,000 genomes project) identified a homozygous *TRAPPC1* splice variant (Chr17:g.7931501C>T, NM_021210.5: c.170+5G>A; p.? [hg38], absent in gnomAD) as a possible cause of her disorder (Figure A3.1, Table A3.3). Splicing prediction tools (splice AI, MaxEnt and Human Splice Finder (HSF)) give mixed predictions on splicing (Table A3.3) and this variant is classified as a VUS according to ACMG/ACGS criteria [402]. Other variants identified by WES/WGS that were unable to be excluded as the cause of disease are listed in Table A3.4.

CASES	Variant	GRCh38 (hg38):	gnomAD v2.1.1 HET	gnomAD v3.1.2 HET	gnomAD HOM	gnomAD All AF	SIFT	Provean	Polyphen-2	REVEL	Splice AI	ACMG Class
TRAPPC1: NM_021210.5												
Family 1	c.362_363del, p.(Val121Alafs*3)	Chr17:g.7930680GCA>G	47	28	0	0.0001662	*NA*					Pathogenic
Family 1	c.64_72del, p.(22_24del)	Chr17:g7931758_7931766del	0	0	0	0	*NA*					Likely pathogenic
Family 2	c.293A>C, p.(His98Pro)	Chr17:g.7931027T>G	0	0	0	0	0.035	-2.63	0.921	0.725	0.02	VUS
Family 2	c.215A>G, p.(His72Arg)	Chr17:g.7931105T>C	1	1	0	0.000000658 4	0.004	-5.43	0.983	0.804	0.01	VUS
Family 3	c.170+5G>A, p.?	Chr17:g.7931501C>T	0	0	0	0	*NA*				0.29	VUS
TRAPPC5: NM_174894.2												
Family 1	c.416delG, p.(Cys139Serfs*33)	Chr19:g.7682668TG>T	3	4	0	0.00002627	*NA*					Pathogenic
Family 1	c.296G>T, p.(Gly99Val)	Chr19:g.7682549G>T	26	61	0	0.0004007	0.000	-8.80	0.999	0.655	0.00	Likely pathogenic
TRAPPC8: NM_014939.5												
Family 1	c.2935_2937del p.(Ser979del)	Chr18:g.31857790CACT>C	0	0	0	0	*NA*					VUS
Family 2	c.4019A>G, p.(Asn1340Ser)	Chr18:g.31832138T>C	25	19	0	0.0001250	0.069	-3.13	0.960	0.286	0.44	VUS
Family 3	c.1624T>G p.(Leu542Val)	Chr18:g.31890839A>C	0	0	0	0	0.000	-2.59	1.0	0.747	0.00	VUS
Family 3	c.91C>T p.(Leu31Phe)	Chr18:g.31942674G>A	0	0	0	0	0.053	-1.90	0.183	0.099	0.00	VUS

Table A3.3: Biallelic *TRAPPC1*, *TRAPPC5* and *TRAPPC8* variants identified in affected individuals from seven families

Includes data from gnomAD (v2.1.1/v3.1.2) of allele count and variant frequencies, pathogenicity prediction tools and variant classification according to ACMG and UK ACGS guidelines [401, 402]. Abbreviations: HOM; homozygous, HET; heterozygous, AF; allele frequency, ACMG; American College of Medical Genetics, VUS; variant of uncertain significance.

Family	Gene	Zygosity	GRChr38:g	c.Nomen	p.Nomen	gnomAD v2.1.1 MAF	gnomAD v3.1.2 MAF	*In silico* predictions		
								Polyphen	SIFT	REVEL
TRAPPC1										
3	ZNF347	HOM	Chr19:g.53149242C>CA	NM_032584.2: c.141G>GT	NM_032584.2: p.(Leu47Leufs*)	-	-	N/A	N/A	N/A
	KCNK9	HET (mat)	Chr8:g.139618460C>T	NM_001282534.2: c.923G>A	NM_001282534.2: p.(Gly308Asp)	0.0001672	0.00003942	Benign	Tolerated	0.062
	RGS21	HOM	Chr1:g.192347328G>C	NM_001039152.3: c.27G>C	NM_001039152.3: p.(Arg9Ser)	-	-	Benign	Tolerated	0.128
	TTN	HET (pat)	Chr2:g.178618350G>A	NM_001267550.2: c.47108C>T	NM_001267550.2: p.(Thr15703Ile)	0.000008069	0.000006587	Probably damaging	Damaging	0.425
	TTN	HET (mat)	Chr2:g.178730942G>A	NM_001267550.2: c.17723C>T	NM_001267550.2: p.(Ala5908Val)	-	-	Benign	Tolerated	0.328
	KIR3DL1	HOM	Chr19:g.54830217T>C	NM_013289.3: c.1277T>C	NM_013289.3: p.(Ile426Thr)	-	-	Benign	Tolerated	0.031
	KIR3DL1	HOM	Chr19:g.54830217T>A	NM_013289.3: c.1279T>A	NM_013289.3: p.(Leu427Met)	0.000008461	0.000007382	Benign	Tolerated	0.033
	KIR3DL1	HOM	Chr19:g.54830258G>A	NM_013289.3: c.1318G>A	NM_013289.3: p.(Val440Ile)	0.000004234	-	Possibly damaging	Tolerated	0.076
	SLC4A3	HET (mat)	Chr2:g.219632295C>T	NM_005070.4: c.994C>T	NM_005070.4: p.(Arg332Cys)	0.0001062	0.00005255	Possibly damaging	Damaging	0.412
	SLC4A3	HET (pat)	Chr2:g.219629375C>T	NM_005070.4: c.449C>G	NM_005070.4: p.(Ala150Val)	0.000004320	-	Benign	Tolerated	0.022
	ITM2C	HET (pat)	Chr2:g.230865060G>T	NM_030926.6: c.35G>T	NM_030926.6: p.(Gly12Val)	-	-	Benign	Tolerated	0.045

	Gene	Zygosity	Genomic	Transcript (c.)	Protein (p.)			PolyPhen	SIFT	Score
	ITM2C	HET (mat)	Chr2:g.230878473C>T	NM_030926.6: c.*374=	NM_030926.6: c.*374=	0.00003192	0.00004638	N/A	N/A	N/A
	SGTA	HET (*de novo*)	Chr19:2759275T>C	NM_003021.3: c.719A>G	NM_003021.3: p.(Asn240Ser)	-	-	Benign	Tolerated	0.119
TRAPPC3										
2	*CDC6*	HOM	Chr17:g.40293571C>A	NM_001254.3: c.776C>A	NM_001254.3: p.(Ala259Asp)	0.00001988	0.000006572	Benign	Deleterious	0.099
TRAPPC8										
1	*PC*	HOM	Chr11:g.66870367C>T	NM_000920.4: c.838G>A	NM_000920.4: p.(Ala280Thr)	0.00006402	0.00003287	Probably damaging	Deleterious	0.860
	CAD	HOM	Chr2:g.27218023A>C	NM_004341.5: c.222+7A>C	NM_004341.5: p.?	-	-	N/A	N/A	Splice AI 0.00
	BCOR	HEMI	ChrX:g.40072410C>T	NM_001123385.1: c.2936G>A	NM_001123385.1: p.(Cys979Tyr)	-	-	Probably damaging	Deleterious	0.378
2	*CLPP*	HOM	Chr19:g.6361595A>AG	NM_006012.2: c.28del	NM_006012.2: p.(Ala10GlyfsTer65)	0.0001474	-	N/A	N/A	N/A
	ADGRV1	HOM	Chr5:g.90807658C>T	NM_032119.4: c.14893C>T	NM_032119.4: p.(Pro4965Ser)	0.000008031	-	Possibly damaging	Deleterious	0.098
	ELP2	HOM	Chr18:g.36159975C>T	NM_018255.4: c.1648C>T	NM_018255.4: p.(His550Tyr)	0.00006370	0.00002628	Benign	Deleterious	0.256
3	*GRIN2A*	HET (*de novo*)	Chr16:g.9763442G>A	NM_001134407.1: c.4102C>T	NM_001134407.1: p.(Leu1368Phe)	-	-	Benign	Deleterious	0.173
	RPL10	HEMI	ChrX:g.154399394C>T	NM_006013.3: c.80C>T	NM_006013.3: p.(Pro27Leu)	-	-	Benign	Deleterious	0.823
TRAPPC14										
2	*OBSL1*	HOM	Chr2:g.219556176G>A	NM_015311.2: c.4453C>T	NM_015311.2: p.(Arg1485*)	0.00002045	0.00002628	N/A	N/A	N/A
	SZT2	HOM	Chr1:g.43427101C>G	NM_015284.3: c.3184C>G	NM_015284.3: p.(Pro1062Ala)	0.0006293	0.0002431	Benign	Tolerated	0.081

3	ZNF433	HOM	Chr19:g.12015231C>T	NM_001080411: c.1636G>A	NM_001080411: p.(Ala546Thr)	0.00003203	0.00001328	Benign	Tolerated	0.021
	TGM1	HOM	Chr14:g.24255088C>G	NM_000359.3: c.1811G>C	NM_000359.3: p.(Ser604Thr)	-	-	Benign	Tolerated	0.127

Table A3.4: Genomic variants identified by WES/WGS in each family that could not be excluded

Abbreviations: HOM, homozygous; HET, heterozygous; HEMI, hemizygous; mat, maternally inherited; pat, paternally inherited; MAF, mean allele frequency; N/A, not applicable.

TRAPPC5

TRAPPC5 is a highly evolutionarily conserved 188 amino acid protein within the core TRAPP complex, that interacts with TRAPPC2, TRAPPC3/L and TRAPPC4 subunits [265, 485]. TRAPPC5 contains a single TRAPP interacting domain [486] and is proposed to localise TRAPP complexes to the Golgi [487]. TRAPPC5 is widely expressed in most tissues (GTex, Broad Institute 2022).

Family 1 comprises a male individual from the USA with biallelic variants in *TRAPPC5* identified as the likely cause of a neurodegenerative disorder characterised by microcephaly, severe GDD, seizures, choreoathetosis, cerebral visual impairment/blindness and neurogenic bladder. Compound heterozygous variants (Chr19:g.7682668TG>T, c.416delG; p.(Cys139Serfs*33) and Chr19:g.7682549G>T, c.296G>T; p.(Gly99Val) [hg38]) in *TRAPPC5* (NM_001042462.2) were identified by trio WES. The p.(Gly99Val) variant is located within the TRAPP domain, with low AF 0.0004007 in gnomAD and no homozygotes, amino acid conservation to zebrafish and multiple *in silico* tools predicting pathogenicity (Figure A3.2, Table 3.3). The p.(Cys139Serfs*33) variant is one of only five LOF variants listed in gnomAD v2.1.1/v3.1.2, none of which are homozygous.
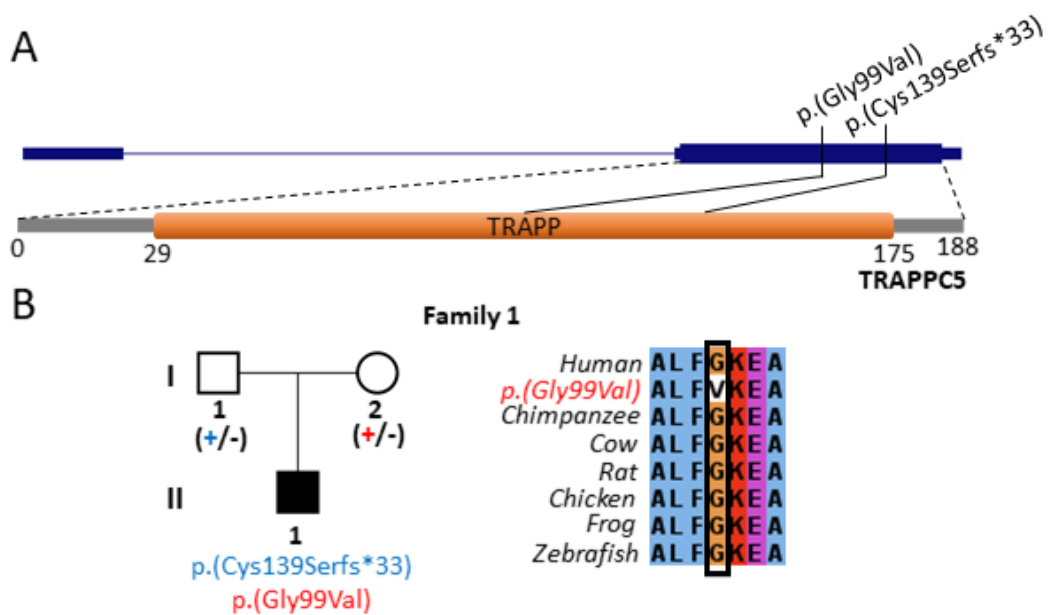
Figure A3.2: *TRAPPC5* variants and family pedigree

(**A**) Schematic of the *TRAPPC5* gene (blue) showing a single coding exon as a wide blue bar, non-coding exonic regions as narrower blue bars and a single intron as a blue line. The TRAPPC5 protein product is shown below including a known TRAPP domain (orange) [341]. The location of *TRAPPC5* variants are indicated by black lines. (**B**) Pedigree of a single family with an affected male individual with segregating biallelic *TRAPPC5* variants '+' indicates the presence of a variant and '-' indicates absence. Multi-species alignment is shown for the p.(Gly99Val) variant within the TRAPP domain with high conservation of the surrounding region.

TRAPPC8

TRAPPC8 is a 1435 amino acid TRAPP III complex subunit protein [265] involved in ER to Golgi transport, and has been identified to interact with the TRAPP III complex via interaction with TRAPPC2 [488], and also interacts with TRAPPC2L and TRAPPC13 subunits [489]. TRAPPC8 has also been proposed to interact with TBC1D14 to regulate the initiation of autophagy [490]. TRAPPC8 contains an N-terminus α-solenoid containing tetratricopeptide repeats (TPR) domain and two C-terminus ASPM, SPD-2, Hydin (ASH) domains [491] (Figure A3.3). TRAPPC8 localises to the centrosome/cilia through ASH domain targeting [491] and has been proposed to play a role in ciliogenesis through interaction with OFD1 [270]. TRAPPC8 is expressed in a wide range of tissues (GTex, Broad Institute 2022).

Three families with individuals affected by a similar neurodevelopmental disorder and biallelic variants in *TRAPPC8* were identified. Families 1 and 2 were identified through the 100,000 genomes project (Figure A3.3). Family 1 comprises a 14 year old male individual from a consanguineous family with severe ID, GDD, febrile seizures, astigmatism, myopia, retinal dystrophy, mild conductive hearing impairment, prominent ear helix and pectus carinatum (Table A3.5). WGS identified a segregating homozygous *TRAPPC8* inframe deletion (Chr18:g.31857790CACT>C, NM_014939.5: c.2935_2937del; p.Ser979del [hg38], absent in gnomAD) (Figure A3.3, Table A3.3). An 18 year old female (Individual II:1) from a consanguineous family (Family 2) has similar clinical features to those observed in Family 1 with additional features of spastic quadriplegia, hyperreflexia, bilateral sensorineural hearing loss, hepatomegaly and scoliosis (Table A3.5). WGS identified a homozygous *TRAPPC8* missense variant (Chr18:g.31832138T>C, NM_014939.5: c.4019A>G; p.(Asn1340Ser) [hg38], gnomAD AF 0.0001250 with no homozygotes) (Table A3.3).
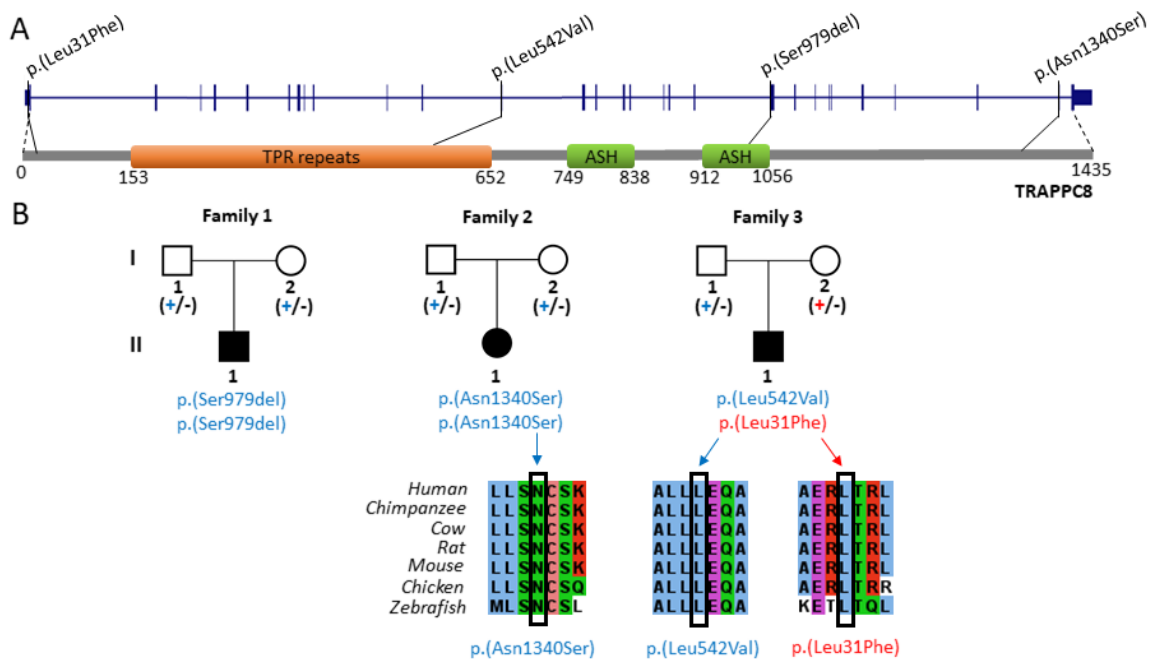
Figure A3.3: *TRAPPC8* variants and pedigrees of affected individuals

(**A**) Schematic of the *TRAPPC8* gene (blue) showing coding exons in wide blue bars, non-coding exonic regions as narrower blue bars and introns as a blue line. The TRAPPC8 protein product is shown below including a known TPR repeat domain (orange) and two ASPM, SPD-2, Hydin (ASH) domains [341]. The location of *TRAPPC8* variants are indicated by black lines. (**B**) Pedigrees of three singleton affected individuals with biallelic *TRAPPC8* variants segregating for an autosomal recessive disorder; '+' indicates the presence of a variant and '-' indicates absence. Multiple species alignments below indicate the conservation of all missense variants across species.

Family 3 was identified from the DDD CAP#303 and comprises a 13 year old affected male (II:1) with features of intellectual disability, GDD, microcephaly, short stature, truncal obesity and Achilles tendon contracture (Table A3.5). Trio WES identified compound heterozygous missense variants in *TRAPPC8* NM_014939.5 (Chr18:g.31890839A>C, c.1624T>G; p.(Leu542Val) and Chr18:g.31942674G>A, c.91C>T; p.(Leu31Phe) [hg38]), both absent in gnomAD (Table A3.3). No other candidate pathogenic variants suspected to be the cause of disease were identified in affected individuals from any of the three families and variants that were unable to be excluded are listed in Table A3.4.

TRAPPC8 shows significant constraint for LOF variants in gnomAD with a pLI score of 1.0, observed/expected LoF=0.07 and LOUEF score of 0.14, although constraint of missense variants is non-significant. There are 22 LOF *TRAPPC8* variants in gnomAD, none of which are listed in homozygous state.

FAMILY	FAMILY 1 II:1	FAMILY 2 II:1	FAMILY 3 II:1
Ethnicity	UK	UK Asian	UK
Genotype	p.(Ser979del)/p.(Ser979del)	p.(Asn1340Ser)/p.(Asn1340Ser)	p.(Leu542Val)/p.(Leu31Phe)
Age at assessment (years)	14	18	13
Sex	M	F	M
GROWTH PARAMETERS			
Birth gestation (weeks)	NK	NK	41
Birth weight kg (SD)	NK	NK	3.3 (-0.53)
Birth OFC cm (SD)	NK	NK	NK
Height cm (SD)	NK	Short stature	136 (-3.25)
Weight kg (SD)	NK	NK	46 (-0.42)
OFC cm (SD)	Microcephaly	Microcephaly	52.6 (-1.98)
DEVELOPMENT			
Intellectual disability	+	+	+
Global developmental delay	+	+ developmental regression	+
Speech impairment	+	+	+ first words 2.5-3y
Motor milestones	Delay	Delay	Delay, sitting 8m, walked 22m
Vision	Astigmatism, myopia, retinal dystrophy	NK	NK
Hearing	Mild conductive impairment	Bilateral SNHL	NK
NEUROLOGICAL			
Tone	NK	NK	NK
Muscle weakness	NK	+ spastic quadriplegia	NK
Spasticity	NK	+ hyperreflexia	NK
Cerebellar signs	NK	NK	NK
Seizures	Febrile seizures	NK	NK
Neuroimaging	NK	Leukodystrophy	NP
EEG	NK	NK	NP
Behavioural abnormalities	NK	NK	Restrictive behaviour
DYSMORPHIC FEATURES			

Facial features	Prominent ear helix	NK	Abnormally folded helix, brachycephaly, submucous cleft hard palate
Limbs	NK	Knee flexion contracture, hip abnormality	Achilles tendon contracture, pes planus
OTHER	Pectus carinatum	Vitiligo, gastrooesophageal reflux requiring gastrostomy, hepatomegaly, scoliosis, vitamin D deficiency	Truncal obesity, feeding difficulties – submucous cleft

Table A3.5: Clinical features of individuals with biallelic variants in *TRAPPC8*

'+' indicates the presence and '-' indicates absence of a feature. Abbreviations: SD; standard deviation, OFC; occipitofrontal circumference, EEG; electroencephalogram, SNHL; sensorineural hearing loss, y; years, m; months, NK; not known, NP; not performed.

A3.3.3 Consolidation of candidate TRAPPopathy genes

TRAPPC3

TRAPPC3 is a core TRAPP complex subunit consisting of 180 amino acids, interestingly each core TRAPP complex consists of two TRAPPC3 (or TRAPPC3L) subunits (Figure 3.1) [105]. Within the TRAPP complex, TRAPPC3 plays a key role in intracellular trafficking through tethering of transport vesicles to the cis-Golgi membrane [492] and has also been identified to play a role in ciliogenesis in retinal pigment epithelium (RPE) cells [269, 491] where it has been identified to interact with TRAPPC9, TRAPPC10 and other proteins implicated in the pathomolecular basis of Bardet-Biedel syndrome (the BBSome). Ciliary membrane biogenesis was impaired through the inhibition of the Rab8 guanine nucleotide exchange factor Rabin8 recruitment to the centrosome [269, 493]. TRAPPC3 is widely expressed, including high levels of expression in the brain (GTex, Broad Institute 2022).

To date only a single family (Family 1) has been published with biallelic segregating variants in *TRAPPC3* in a paper identifying a number of candidate novel ciliopathy genes [281]. Family 1 comprises two affected sisters both with a Bardet-Biedel syndrome (BBS)-like phenotype characterised by retinitis pigmentosa, polydactyly, obesity, developmental delay and intellectual disability, with additional features of dextrocardia and complex congenital heart disease in one sister (Table A3.6, Figure A3.4). This study by Shaheen *et al* (2016) used a combination of autozygosity mapping with exome sequencing to identify a single candidate homozygous missense variant in *TRAPPC3* (Chr1:g.36139776G>A NM_014408.5: c.184C>T; (p.Arg62Trp) [hg38]) as the likely cause of this disorder (Table A3.7).

Families 2 and 3 were identified through the DDD CAP#303. Family 2 comprises a 4 year old male from a consanguineous nuclear family, who was diagnosed with Artemis-related severe combined immunodeficiency (SCID) in the first year of life and found to have a homozygous deletion of exons 1-3 of *DCLRE1C* on microarray CGH, a known cause of SCID and received a successful bone marrow transplant. This individual has several features consistent with a BBS-like condition including severe microcephaly, short stature, obesity, severe intellectual disability and developmental delay, postaxial polydactyly and cerebellar atrophy on MRI neuroimaging (Table A3.6). Trio WES identified a homozygous *TRAPPC3* missense variant (Chr1:g.

36139808A>G, NM_014408.5: c.152T>C; p.(Ile51Thr)) with an allele frequency of 0.000006571 in gnomAD and predicted to be deleterious by multiple *in silico* missense prediction tools (Table A3.7).

Family 3 is a consanguineous family with an 8 year old affected male individual with BBS-like features overlapping those seen in Families 1 and 2 (Table A3.6). MRI neuroimaging was normal, with negative testing for Prader-Willi syndrome. A homozygous *TRAPPC3* variant (Chr1:g.36137826C>A NM_014408.5: c.393G>T; p.(Leu131Phe), AF 0.0002760 in gnomAD), was identified by trio WES as a possible cause of this disorder. No other candidate pathogenic variants suspected to be the cause of disease were identified in affected individuals from these families and variants that could not be excluded are listed in Table A3.4.
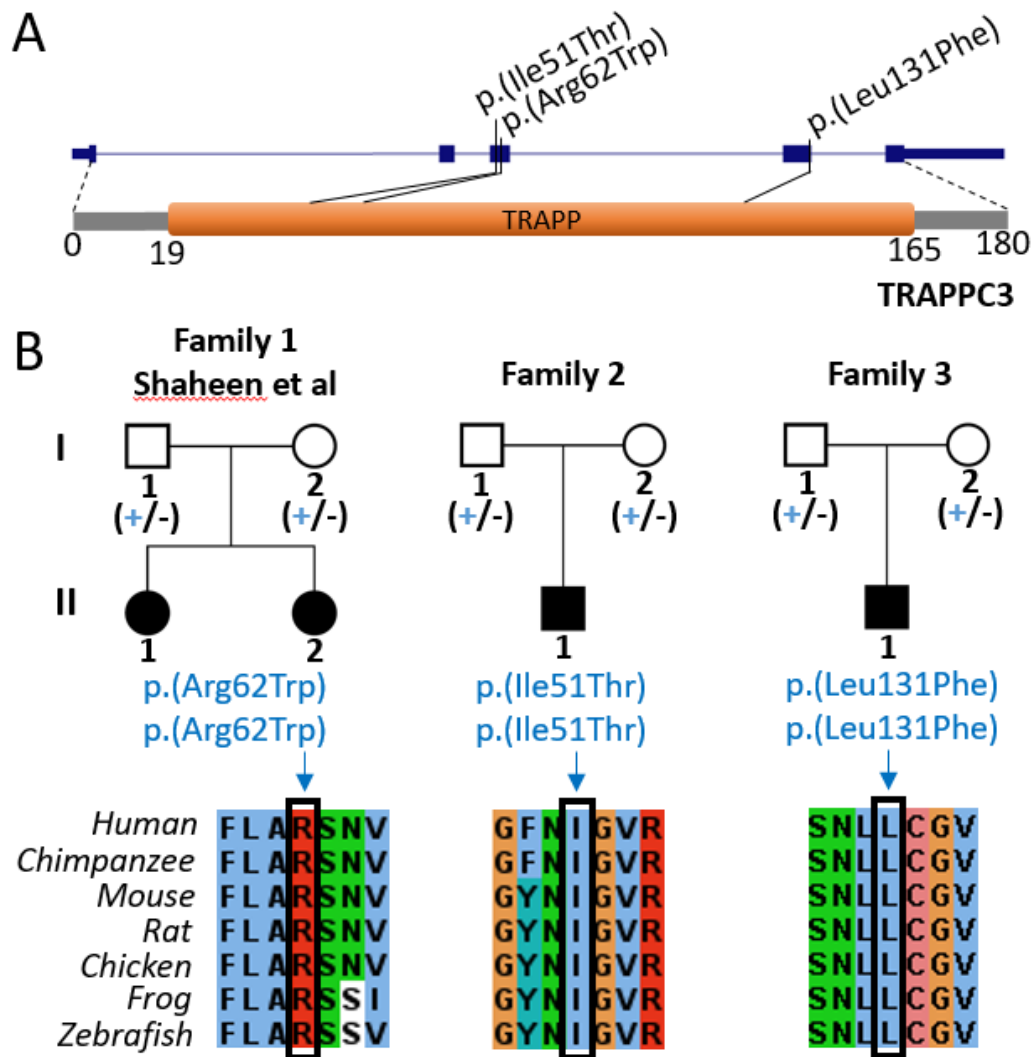
Figure A3.4: *TRAPPC3* variants and pedigrees of affected individuals

(**A**) Schematic of the *TRAPPC3* gene (blue) showing coding exons in wide blue bars, non-coding exonic regions as narrower blue bars and introns as a blue line. The TRAPPC3 protein product is shown below including known TRAPP protein domain (orange) [341]. The location of *TRAPPC3* variants are indicated by black lines. (**B**) Pedigrees of three families of affected individuals with biallelic *TRAPPC3* variants segregating for an autosomal recessive disorder; '+' indicates the presence of a variant and '-' indicates absence. Multiple species alignments below indicate the conservation of all missense variants across species.

FAMILY	FAMILY 1 (Shaheen *et al*)		FAMILY 2	FAMILY 3
Ethnicity	Saudi Arabia	Saudi Arabia	UK	UK
Genotype	p.(Arg62Trp)/p.(Arg62Trp)	p.(Arg62Trp)/p.(Arg62Trp)	p.(Ile51Thr)/p.(Ile51Thr)	p.(Leu131Phe)/p.(Leu131Phe)
Age at assessment (years)	7	16	4	9
Sex	F	F	M	M
GROWTH				
Birth gestation (weeks)	Full-term	Full-term	40	38
Birth weight kg (SD)	NK	NK	2.95 (-1.27)	2.5 (-1.45)
Birth OFC cm (SD)	NK	NK	NK	NK
Height cm (SD)	NK	NK	94.8 (-2.82)	138.4 (+1.13)
Weight kg (SD)	NK	NK	30.6 (+3.92)	56.25 (+3.24)
OFC cm (SD)	NK	NK	45.5 (-4.74)	52.7 (-0.89)
DEVELOPMENT				
Intellectual disability	+	+	+ (severe)	+
Global developmental delay	+	+	+	+
Speech impairment	Delay	Delay	Non-verbal	Delay: first word 3-4y
Motor milestones	Delay	Delay	Delay: sitting 9m, walking 2.5y	Delay: sitting 7m, walking 14m
Vision	Nystagmus, poor night vision, RP	Nystagmus, poor night vision, RP	NK	NK
Hearing	NK	NK	NK	NK
NEUROLOGY				
Tone	NK	NK	NK	NK
Muscle weakness	NK	NK	NK	NK
Spasticity	NK	NK	NK	NK
Cerebellar signs	NK	NK	NK	NK
Seizures	NK	NK	NK	NK
Brain imaging	NK	NK	Cerebellar atrophy	Normal
EEG	NK	NK	NK	NK
Behavioural abnormalities	NK	NK	NK	Impaired social interaction
DYSMORPHIC FEATURES				

Facial features	NK	NK	NK	Deeply set eyes, low anterior hairline, thick lip vermilion, upslanted palpebral fissures
Limbs	Polydactyly	Polydactyly	Postaxial polydactyly	
OTHER	Obesity	Obesity Dextrocardia, complex cardiac disease	Obesity SCID: *DCLRE1C* HOM deletion exon 1-3	Truncal obesity Aortic coarctation, hypospadias, cryptorchidism, gynaecomastia

Table A3.6: Clinical features of individuals with biallelic variants in *TRAPPC3*

'+' indicates the presence and '-' indicates absence of a feature. Abbreviations: SD; standard deviation, OFC; occipitofrontal circumference, EEG; electroencephalogram, m; months, y; years, SCID; Severe combined immunodeficiency, RP; retinitis pigmentosa, y; years, m; months, HOM; homozygous, NK; not known, HOM; homozygous.

CASES	Variant	GRCh38 (hg38):	gnomAD v2.1.1 HET	gnomAD v3.1.2 HET	gnomAD HOM	gnomAD All AF	SIFT	Provean	Polyphen -2	REVEL	Splice AI	ACMG Class
TRAPPC3: NM_014408.5												
Family 1 Shaheen et al	c.184C>T; p.(Arg62Trp)	Chr1:g.36139776G>A	5	5	0	0.00003288	2.81	-7.66	0.479	0.519	0.15	Likely pathogenic
Family 2	c.152T>C; p.(Ile51Thr)	Chr1:g.36139808A>G	0	1	0	0.000006571	2.81	-4.58	0.985	0.797	0.02	VUS
Family 3	c.393G>T; p.(Leu131Phe)	Chr1:g.36137826C>A	48	42	0	0.0002760	0.092	-3.04	0.164	0.294	0.00	VUS
TRAPPC14: NM_018275.5												
Family 1 Perez et al	c.613G>T; p.(Glu205*)	Chr7:g.99755280C>T	32	58	0	0.0003810	*NA*					Pathogenic
Family 2	c.1090C>T; p.(Arg364Cys)	Chr7:g.100156536G>A	11	7	0	0.00004599	0.003	-2.16	0.994	0.265	0.04	VUS
Family 3	c.1393G>A; p.(Glu465Lys)	Chr7:g.100155673C>T	0	1	0	0.000006569	0.004	-3.64	0.982	0.331	0.01	VUS

Table A3.7: Biallelic *TRAPPC3* and *TRAPPC14* variants identified in affected individuals from six families

Includes data from gnomAD (v2.1.1/v3.1.2) of allele count and variant frequencies, pathogenicity prediction tools and variant classification according to ACMG and UK ACGS guidelines [401, 402]. Abbreviations: HOM; homozygous, HET; heterozygous, AF; allele frequency, ACMG; American College of Medical Genetics, VUS; variant of uncertain significance.

TRAPPC14

TRAPPC14 (c7orf43/MAP11) has only recently been identified as a possible TRAPP complex protein, forming a complex-specific component of the TRAPP II complex alongside TRAPPC9 and TRAPPC10 [262]. TRAPPC14 plays a role in Rabin8 binding to the TRAPP II complex during ciliogenesis, resulting in preciliary vesicle tethering at the centriole. Although further studies are required to clarify the role of TRAPPC14 in the TRAPP II complex as it is not required for its guanine nucleotide exchange factor activity [263, 264] and loss does not affect complex assembly or stability. TRAPPC14 is widely expressed, including high levels of expression within the cerebellum (GTex, Broad Institute 2022).

Perez *et al* (2019) [320] identified that *TRAPPC14* variants in both humans and zebrafish result in a microcephalic neurodevelopmental disorder phenotype. A single consanguineous family (Family 1 in Figure A3.5 and Table A3.8) with 3 siblings (2 male, 1 female) affected by a neurodevelopmental disorder characterised by intellectual disability, developmental delay, postnatal microcephaly, short stature, ADHD and findings of thin corpus callosum and reduced white matter volume on MRI neuroimaging. A single individual had a tethered spinal cord with a lipoma of filum terminale (Table A3.8). A combination of genome-wide autozygosity mapping of all three affected individuals and WES of a single individual (II:2) identified a single candidate homozygous nonsense variant (Chr7:g.99755280C>T, NM_018275.5: c.613G>T; p.(Glu205*) [hg38]) in *TRAPPC14* (Figure A3.5) located within a 12.4Mb shared region of homozygosity. Perez *et al* showed that the protein product was unlikely to undergo nonsense mediated decay through RT-qPCR studies of lymphocytes from affected and unaffected individuals.
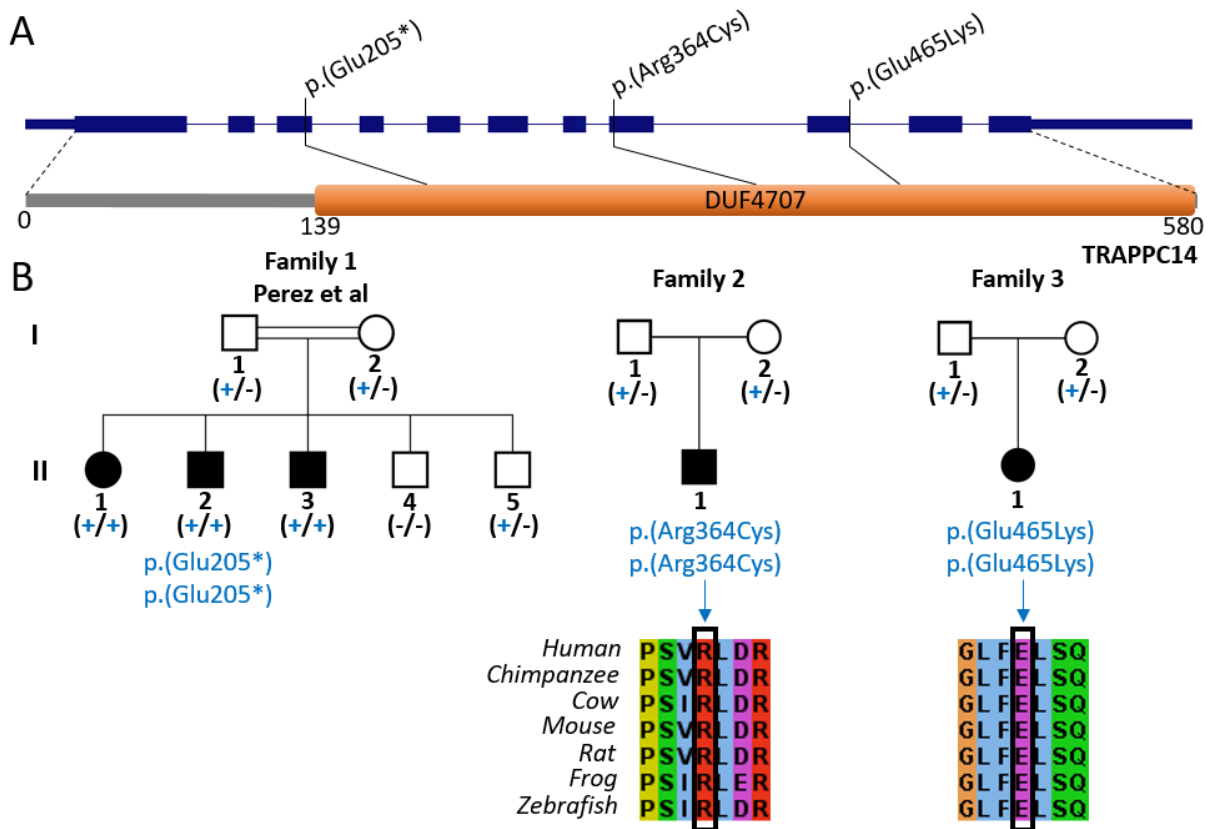
Figure A3.5: *TRAPPC14* variants and pedigrees of affected individuals

(**A**) Schematic of the *TRAPPC14* gene (blue) showing coding exons in wide blue bars, non-coding exonic regions as narrower blue bars and introns as a blue line. The TRAPPC14 protein product is shown below including a domain of unknown function (DUF4707) (orange) [341]. The location of *TRAPPC14* variants are indicated by black lines. (**B**) Pedigrees of three families of affected individuals with biallelic *TRAPPC14* variants segregating for an autosomal recessive disorder; '+' indicates the presence of a variant and '-' indicates absence. Multiple species alignments below indicate the conservation of all missense variants across species.

FAMILY	FAMILY 1 (Perez *et al*)			FAMILY 2	FAMILY 3
Ethnicity	Bedouin Israeli	Bedouin Israeli	Bedouin Israeli	UK	UK
Genotype	p.(Glu205*)/p.(Glu205*)	p.(Glu205*)/p.(Glu205*)	p.(Glu205*)/p.(Glu205*)	p.(Arg364Cys)/ p.(Arg364Cys)	p.(Glu465Lys)/ p.(Glu465Lys)
Age at assessment (years)	6	7	5	13	
Sex	F	M	M	M	F
GROWTH					
Birth gestation (weeks)	40	40	40	NK	NK
Birth weight kg (SD)	2.94 (-0.89)	3.08 (-0.84)	2.6 (-1.8)	NK	NK
Birth OFC cm (SD)	32 (-1.76)	32.5 (-1.87)	32.5 (-1.87)	NK	NK
Height cm (SD)	72 (-3.0)	84 (-0.98)	80 (-2.56)	Short stature	NK
Weight kg (SD)	10.8 (-1.32)	12 (-2.5)	10.8 (-0.11)	NK	NK
OFC cm (SD)	39 (-5.24)	42 (-4.98)	41 (-5.25)	NK	Microcephaly
DEVELOPMENT					
Intellectual disability	+	+	+	+	+
Global developmental delay	+	+	+	+	+
Speech impairment	Delay, few words only	Delay, few words only	Delay, few words only	NK	NK
Motor milestones	Delay	Delay	Delay	Delay	NK
Vision	No concerns	No concerns	No concerns	NK	NK
Hearing	No concerns	No concerns	No concerns	NK	NK
NEUROLOGY					
Tone	NK	NK	NK	NK	NK
Muscle weakness	NK	NK	NK	NK	NK
Spasticity	NK	NK	NK	NK	NK
Cerebellar signs	NK	NK	NK	NK	NK
Seizures	NK	NK	NK	NK	NK
Neuroimaging	Microcephaly, thin corpus callosum, reduced white matter volume	Microcephaly, thin corpus callosum	Microcephaly, thin corpus callosum	NK	NK
EEG	NP	NP	NP	NP	NP
Behavioural abnormalities	ADHD	ADHD	ADHD	NK	NK

DYSMORPHIC FEATURES					
Facial features	-	-	-	NK	NK
Limbs	-	-	-	NK	NK
OTHER	Tethered spinal cord, lipoma of filum terminale	Nil	Nil	NK	NK

Table A3.8: Clinical features of individuals with biallelic variants in *TRAPPC14*

'+' indicates the presence and '-' indicates absence of a feature. Abbreviations: SD; standard deviation, OFC; occipitofrontal circumference, EEG; electroencephalogram, m; months, y; years, ADHD; attention deficit hyperactivity disorder, NK; not known, NP; not performed.

Family 2 comprises a 13 year old male individual identified through the 100,000 genomes project with features of intellectual disability, global developmental delay and short stature. Trio WGS identified a single tier 1 homozygous nonsense *OBSL1* variant (Chr2:g.219556176G>A NM_015311.2: c.4453C>T, p.(Arg1485*)) (Table 3.4) reported as a VUS due to an inconsistent phenotype as the associated condition (3-M syndrome 2, OMIM #612921) [494] only accounts for the feature of short stature in this individual and not neurocognitive impairment, suggesting a possible second diagnosis for this individual. A homozygous *TRAPPC14* variant was identified as a possible cause of intellectual disability and developmental delay in this individual (Chr7:g.100156536G>A, NM_018275.5: c.1090C>T; p.(Arg364Cys)) (Table A3.7).

Family 3 was identified through the Decipher database (v11.10, https://www.deciphergenomics.org/) and includes a single female individual with features of microcephaly, intellectual disability and developmental delay (Table A3.8). WES identified a highly conserved homozygous *TRAPPC14* missense variant (Chr7:g.100155673C>T, NM_018275.5: c.1393G>A; p.(Glu465Lys)) as a possible cause of this individuals phenotype (Figure A3.5). No other candidate pathogenic variants suspected to be the cause of disease were identified in affected individuals from these families and variants that could not be excluded are listed in Table A3.4.

A3.4 Discussion

Two TRAPP complexes have been identified in humans, TRAPP II and TRAPP III, which share the same eight core TRAPP complex proteins, with eight further complex-specific subunits (Figure 3.1). To date 11 TRAPP complex subunits have been associated with human disease (Table 3.1) and this study describes the identification of three further potential TRAPPopathies associated with variants in *TRAPPC1*, *TRAPPC5* and *TRAPPC8* subunits, and consolidation of two other candidate TRAPPopathy disorders (TRAPPC3 and TRAPPC14) (Table A3.1).

TRAPPC1-related disorder was identified in three candidate families in which affected individuals display similar features of intellectual disability, developmental delay with prominent motor delay; all are non-ambulatory and a single individual has myopathy. The TRAPPC1 protein has significant homology with TRAPPC4 and has a synbindin-like domain (Figure A3.1) [486]. TRAPPC4/synbindin is highly expressed in differentiating neurons and interacts with syndecans that act as receptors involved in intracellular communication and recycling endosomes. Syndecans are found on dendritic surfaces and are proposed to facilitate dendritic spine formation through interaction with TRAPPC4/synbindin [329]. The homology of TRAPPC1 and TRAPPC4 proteins suggests that TRAPPC1 is likely to also interact with syndecans and play a role in membrane trafficking and dendritic spine formation. Indeed, there are similarities between neurodevelopmental disorders caused by variants in *TRAPPC1* and *TRAPPC4* (Table A3.1) likely reflecting an overlapping mechanism of pathogenicity, although further studies will be required to examine this further. Cell and molecular studies to evaluate potential functional impairment associated with the *TRAPPC1* variants presented here are currently underway, performed by Michael Sacher's laboratory (Concordia University).

A single family with one affected individual with a neurodevelopmental disorder suspected to be caused by pathogenic/likely pathogenic compound heterozygous variants in *TRAPPC5* was also identified. While these studies are not sufficiently conclusive to describe *TRAPPC5* as a novel disease gene for a further TRAPPopathy disorder, these findings may ultimately be confirmed and aid definition of this condition. In the future, a complete list of other variants identified by WES should be reviewed for this individual to confirm the *TRAPPC5* variants as the most likely candidate cause

of their disorder. The identification of further individuals with similar phenotypes and biallelic variants in *TRAPPC5* should also be sought, alongside functional confirmation of pathogenicity of these variants.

A *TRAPPC8*-related disorder was potentially identified in three singleton individuals with biallelic variants in *TRAPPC8*. While all although all variants are currently classed as VUS, they represent a potential genetic explanation for further studies (each affected individual has other candidate variants that are unable to be excluded; Table A3.4). These individuals have overlapping clinical features including microcephaly, intellectual disability, developmental delay and joint contractures, with retinal dystrophy in one individual and truncal obesity in another. In the future further phenotypic characterisation should be performed to assess the degree of similarity of phenotypes, and whether any further ciliopathic features are present; if concordant these variants may represent a common genetic aetiology. Further clinical, genetic and functional studies are also required to confirm this as a neurodevelopmental disorder gene. Interestingly TRAPPC8 has been identified as a multifunctional protein with roles in ciliogenesis [270], centrosomal functions [491] and autophagy [490], as well as TRAPP III complex function in membrane trafficking and this could give rise to differing phenotypes in individuals with *TRAPPC8*-related disorder.

These studies also provide further genetic data to support and consolidate previous studies identifying biallelic variants in *TRAPPC3* and *TRAPPC14* genes as causative of distinct TRAPPopathy disorders. Two further unrelated individuals with a strikingly similar BBS-like phenotype to that of the family described by Shaheen *et al* [281] were identified. While both have homozygous *TRAPPC3* missense variants classed as VUS, future functional studies will clarify pathogenicity, and the identification of further individuals with the disorder will also corroborate these findings. In view of the known association of TRAPPC3 with ciliogenesis this provides a potential pathogenic mechanism underlying these observed features [269, 491]. Additionally, two individuals were identified with possible *TRAPPC14* variants as causative of a neurodevelopmental disorder, although further clinical characterisation is again required to confirm phenotypic overlap with the family previously published by Perez *et al* [320], as well as functional studies to confirm pathogenicity of these variants.

Overall TRAPPopathy disorders appear to involve four broad phenotypic groups, often with overlapping neurocognitive features (Table A3.1). This includes a purely skeletal dysplasia phenotype (TRAPPC2), Golgipathies with features of microcephaly, intellectual disability, developmental delay, seizures and white matter brain abnormalities (TRAPPC1, TRAPPC6B, TRAPPC9, TRAPPC10 and TRAPPC14); ciliopathy disorders with additional BBS-like phenotype of polydactyly, RP, obesity and cardiac anomalies (TRAPPC3, TRAPPC8 and TRAPPC12); and myopathy/limb girdle muscular dystrophy with features of myopathy, raised CK, movement disorders like dystonia, chorea, tremor (TRAPPC2L, TRAPPC5 and TRAPPC11). With the exception of TRAPPC2, there therefore appears to be a good degree of clinical overlap between these groups of TRAPPopathies.

Some well characterised TRAPPopathy disorders have been identified to have distinct phenotypes caused by variants within the same TRAPP subunit. For example variants in *TRAPPC11* have been associated with predominantly myopathic or muscular dystrophy phenotypes with raised CK and liver abnormalities. Some have additionally been associated with microcephaly and seizures. This may be related to the multiple functions identified to be associated with TRAPPC11 [105], including membrane trafficking [265, 495, 496], regulation of glycosylation of proteins/lipids [273, 315] and autophagy [497]. A number of TRAPP complex proteins have been identified to play a role in ciliogenesis, including TRAPPC3 [269, 491], TRAPPC8, TRAPPC12 [270], TRAPPC14 [262] and TRAPPC10 (personal communication, Michael Sacher). Others have been identified to play a role in centrosomal/microtubule function (TRAPPC8, TRAPPC12) [491, 498] and autophagy (TRAPPC4, TRAPPC8, TRAPPC11 and TRAPPC13) [282, 490, 497, 499, 500]. A picture is emerging that a core phenotype of an autosomal recessive microcephalic neurodevelopmental phenotype may be associated with the disruption of TRAPP complex function in membrane trafficking within the secretory pathway, with additional phenotypic features resulting from disruption of additional functions of specific TRAPP complex proteins either in isolation or with other complex members in cilia function, centrosomal processes, autophagy and glycosylation. Further studies are required to examine these hypotheses in detail. Overall, these studies highlight the complexity of studying autosomal recessive neurodevelopmental disorders that are often genetically and phenotypically heterogeneous with multiple interacting partners and pathways.

4 APPENDIX: Methods and results of mouse and cell studies of *KPTN*-related disorder

A4.1 Methods

A4.1.1 Genetics studies

Clinical case identification

Affected individuals were recruited with informed consent through an existing clinical-genetic programme to define the spectrum and causes of genetic disease amongst the Amish community (see www.WOHproject.com), as well as through contact from responsible treating clinicians following diagnostic or research-based exome sequencing submitted to GeneMatcher [244], the DDD project (CAP#255) [242], the 100,000 genomes project [501], and through direct contact from families (including from the KPTN Alliance). Individuals 1-4 were identified within the Ohio Amish community, and are distantly related to the extended family originally reported (Figure 4.1, Family 1). Individuals 5 and 6, two siblings from Michigan (USA), were identified by contact through GeneMatcher [244], four individuals were identified by personal correspondence from the patients family from Virginia, (USA) (Patient 7), New York (Patient 18) and Canada (siblings 20 and 21), six individuals were identified through direct contact from their treating clinicians from Denmark (Patient 8), the Netherlands (siblings 9 and 10, Patient 11 and Patient 14), London (UK, Patient 17), and Liverpool (UK, Patient 19), with two cases from Leicester (UK, siblings 15 and 16) identified through DDD CAP study #255 and two cases from Newcastle (UK, siblings 12 and 13) identified through the 100,000 genomes project.

Founder variant haplotype analysis

Single-nucleotide polymorphism (SNP) genotyping, using Illumina HumanCytoSNP-12 2.1 Beadchip array, incorporating ~330,000 SNP markers, was performed for eight previously published individuals [333] (three homozygous for the p.(Ser259*) variant (Individuals IX:4, IX:5 and IX:6), five compound heterozygous [p.(Ser259*)| p.(Met241_Gln246dup)] (Individuals IX:8, IX:10, IX:11, X:1 and X:2)) and four newly identified affected individuals (three homozygous for the p.(Met241_Gln246dup) variant (Individuals 7, 9 and 10) alongside one individual compound heterozygous with

a different variant (Individual 14)). Haplotype analysis was manually performed to identify SNPs flanking both Amish community founder gene variants, which enabled us to define a core, shared haplotype length for each variant. From this we estimated haplotype age using the Gamma method proposed by Gandolfo *et al* (2014) (https://shiny.wehi.edu.au/rafehi.h/mutation-dating/) [502].

A4.1.2 Mouse studies

Mouse studies were performed by Maria Levitin, Sebastian Gerety and Binnaz Yelcin and form part of the studies for PhD thesis by Maria Levitin [382].

Mouse production

The mouse model was generated at the Wellcome Sanger Institute. The Kptn[tm1a(EUCOMM)Wtsi] mice were kept on a C57BL/6NTac background (Taconic Biosciences). The tm1a 'knockout-first' allele was generated by the insertion of an IRES:lacZ trapping cassette and a LoxP flanked promoter-driven neo cassette into an intron, disrupting *Kptn* gene function at the mRNA level by interfering with transcription downstream of the cassette site [380, 479, 503] (Figure A4.1).

Housing and breeding of mice and experimental procedures were carried out under the authority of UK Home Office project and personal licenses. All procedures were carried out in accordance with the Animal Welfare and Ethical Review Body of the Wellcome Sanger Institute and the UK Home Office Animals (Scientific Procedures) Act of 1986, under UK Home Office PPL 80/2472 and PPL P6320B89B. Mice were housed in pathogen-free mouse facilities with 12-h light/dark cycle. Food and water were available *ad libitum*, except where indicated otherwise.

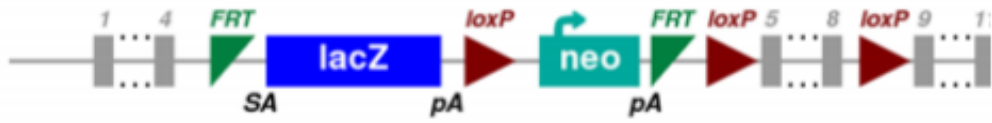**Figure A4.1: Construct used to generate the Kptn$^{-/-}$ (Kptn$^{Tm1a/Tm1a}$) mice**

Main insertion cassette is placed between exons 4 and 5, with a further LoxP between exons 8 and 9. The endogenous allele is disrupted by the insertion and generates a frame shift followed by a premature stop codon. *Kptn* expression is ablated by the insertion. Data courtesy of Dr Maria Levitin and Dr Sebastian Gerety.

Mouse behavioural paradigms

In all assays, the mice were tracked using over-head infrared video cameras and a user-independent automated video tracking software Ethovision XT 8.5 (Noldus Information Technology, The Netherlands). All experiments were carried out on mice aged 10-22 weeks. The mice were handled for 2-3 days before the onset of testing. All mice were habituated to the behavioural room in their home cages for ≥ 30 minutes under the same light condition as the test.

Open field assay

To test for locomotor capabilities, open field test was run as previously described [381]. A period of movement was defined when the mouse reached a velocity of 2 cm/s over two frames; a period of non-movement was defined when velocity was lower than 1.75 cm/s over two frames. In all open field trials, mice were placed in the left corner of the open field arena. Two animals were tested in parallel in two separate open fields (74 cm x 74 cm), for a total observation time of 10 minutes. A centre zone was designated with equidistant borders to the open field walls (8 cm). The arenas were cleaned thoroughly with ethanol free wipes to remove odour traces.

Light/dark box assay

Light/dark box assay was used to test for anxiety-related behaviours, as previously described [383, 384]. The assay was conducted in a darkened room. After 30 minutes of habituation to the room in their home cages, mice were separated into new cages (with clean bedding, food pellets, fun tunnel, and some used bedding), with two mice per cage. The light/dark box was placed into an open field arena, positioned in the same location in the open field arena in all experiments. The light zone had two spotlights facing down directly into the compartment, illuminating the whole of the 'light' compartment evenly (300-400 Lux). The dark zone was sheltered from the light by an opaque lid and a door at the opening between the two compartments. The mouse was placed in the centre of the dark compartment. The door was released at the start of the trial. Trial duration was 10 minutes, and two mice were tested in parallel. The time spent in each of the compartments, as well as the frequency of the transition between zones, was recorded. Preference score was calculated as the time spent in one of the compartments out of the total time of the assay.

Social recognition (SoR)

To test olfactory-mediated memory, the SoR assay was used based on previously published protocols [385, 504]. Test animals were habituated to the room (and red light) for half an hour in their home cages and were then habituated to the arenas for 10 minutes before the start of the testing. Stimuli were matched to test animals according to their weight, size, and gender. The familiar-unfamiliar stimuli pair were from different breeding colony, but of the same genetic background (129P2/OlaHsdWtsi + C57BL/6J), and counterbalanced for novelty. In trial 5, C57BL/6NTac mice were used. For each trial, the same stimulus was used for 3 test animals to reduce the number of animals needing to be anaesthetised. Stimulus animals were subject to non-terminal anaesthesia with ketamine/xylazine (intraperitoneal injection 1 g/0.1g per kg of body weight) and were recovered with atipamezole (subcutaneous injection 0.5mg) in a home cage with a warm water bottle at the end of the experiment.

Pairwise discrimination

Pairwise visual discrimination (PD) assay was run using the automated touchscreen technology [505-507]. The protocol used was adapted from previous studies and the Campden instruction manual for PD task (for mouse touch screen systems and ABET II) [505, 506, 508]. Because the assay is dependent on appetitive reward, the mice were food restricted to achieve a gradual reduction of 10-15% of the initial body weight, and this weight was maintained throughout training and testing. The weight of the mice was measured daily, and the food was adjusted accordingly. All the mice were progressed through the assay on an individual basis. Mice were tested in the same testing box each time throughout all the sessions. Once each mouse reached the criteria in the PD task phase they were no longer tested. In order to represent the percentage of correct trials over sessions in PD task, for those mice that completed the PD task earlier, the criteria mean (correct trial percentage over the last two days) was calculated, and this number was included into the overall mean for the subsequent sessions. Therefore, when the percentage of trials was plotted against sessions, each session had an equal number of mice, even though some mice finished the assay faster than others. The actual number of days taken to reach criteria was plotted

separately. There was a predefined 25 days cut-off point, after which if mice did not complete the criteria, they were excluded.

Barnes maze

The Barnes maze assay was run as described previously [388, 509, 510]. External spatial cues of were placed around the walls of the room, to aid navigation. The mice were tested in a darkened room with overhead lights facing the Barnes maze tables, under maximum illumination. Three days before the first session, their cages were cleaned and an escape box was placed in the centre of the clean home cage. Mice were placed into the box and allowed to climb out of it. If the mouse did not leave the box after a minute, it was gently encouraged back out into the cage. The mice were tested sequentially and in the same order on every day, with two mice tested in parallel on two tables. Fifteen minutes before testing, mice were singly housed in new cages with bedding, food, a "fun tunnel", and a handful of old bedding from the home cage of the mouse.

Transcardial perfusions

Mice were terminally anaesthetised by intraperitoneal injection of 0.1mL pentobarbitone sodium. After the mice were no longer responsive, 20 mL of cold PBS was injected into the left ventricle of the exposed heart at low speed, followed by 20 ml of 4% PFA, either by hand (20-50 ml syringe) or using a perfusion pump. After the transcardial perfusion, skulls and brain segments were collected.

MRI (Magnetic Resonance Imaging)

For MRI analysis, skulls were collected after transcardial perfusion, and stored in formalin solution at 4°C. MRI imaging and analysis were performed, using previously published methodology [511]. Brains were scanned using a Bruker PharmaScan 47/16 system at 4.7T with a manufacturer-provided birdcage transmit-receive coil. The images were registered into the same stereotactic space and segmented into images of four tissue types: grey matter, white matter, cerebrospinal fluid and 'other'. This was followed by voxel-based quantification of brain volume [512]. To control the type I error rate due to multiple comparisons, an adjusted p-value was used for a false discovery rate at $p < 0.05$.

Microcomputed tomography

Mouse skulls were collected after transcardial perfusion and stored in formalin solution at 4°C until they were ready to be scanned. High-resolution images were acquired using the 1172 Bruker microCT scanner. The scanning parameters were set to: 2000 x 1332 object position, pixel size 9.89, Medium camera size. Prior to acquisition, we ensured the average intensity value (av=xx) was in the acceptable range of 40-60% (no more than 70%) and the 'min' value was ≥10%. DICOM files generated by the Bruker software were loaded into 3DimViewer (V3.0, http://www.3dim-laboratory.cz) for surface model creation using default thresholding (1482/7000) and saved as an .STL file. The STL files were loaded individually into Meshlab software, and individual landmark locations [390] were recorded (V2016.12, www.meshlab.net). Inter-landmark distances were calculated in excel. Significant differences in mean linear distances were identified using a student's t-test (p<0.05). Visualisation, 3d rendering and resectioning was performed in Meshmixer V3.5.474 (www.meshmixer.com).

Histomorphometric analysis

For the histomorphological analysis, adult mice aged 16 weeks (n=8 per genotype and sex) were transcardially perfused, the brain was removed from the skull and post-fixed in formalin for 48 hours, after which it was kept in PBS. For the P0 brains, male mice were culled at birth using schedule 1, brains were removed from skull and post-fixed for 12 days before being transferred into PBS. The morphometric analyses were using a previously published method [324], examining the same three coronal brain regions (section 1: Bregma +0.98mm; section 2: Bregma -1.34mm; and section 3: Bregma -5.80mm) in all mice. Mice were separately analysed according to their gender. Sections were scanned using Slide scanner (Hamamatsu, NanoZoomer 2.0HT, C9600 series) and accessories (racks and NanoZoomer digital pathology, version 2.5.64 software) at 20x magnification and analysed using ImageJ (78 measurements for p20 and 16 week animals, 53 measurements for P0 animals).

Immunohistochemistry

Adult mice were euthanized by CO_2 asphyxiation followed by thoracotomy. All animals were perfused with ice-cold saline and 4% PFA. Brains were dissected out and post-fixed in 4% PFA for 24hrs. After fixation, brains were embedded in paraffin wax, sliced on a microtome at 8µm, mounted on slides, and hot air dried overnight. Following

drying, brains were deparaffinised, subjected to antigen retrieval via citrate buffer, and washed in distilled water. Washed slides were blocked for 1hr at room temperature (RT) in blocking solution containing 0.3% tween PBS with 5% normal goat serum and incubated in primary antibodies overnight in blocking solution. Primary antibodies used were PS6 (ser240/244; Cell Signalling Technology; 1:1000; mouse anti-rabbit polyclonal) and S6 (NSJ Bioreagents; 1:100, mouse anti-rabbit monoclonal). After incubation with primary antibodies, slides were washed in PBS and placed in blocking solution with secondary antibodies for 2 hr at RT. Secondary antibodies were conjugated with Alexa 594 (Invitrogen; 1:1000; goat anti-rabbit) and 647 (Invitrogen; 1:1000; goat anti-rabbit) fluorophores, respectively. After two hours, slides were washed in PBS and distilled water and finally coverslipped with Vectashield media containing DAPI. Slides were then imaged on a Nikon spinning disk confocal microscope and digital micrographs were created in Imaris software (Oxford Instruments).

Tissue lysis and immunoblotting

For protein extracts used in immunoblotting, perfused tissue was collected as described above, snap frozen in liquid nitrogen, and stored in -80°C until lysis. Tissue was lysed in 1x lysis solution (protease + phosphatase inhibitor cocktail in TPER) in the Qiagen TissueLyser LT with sterile steel beads and operated at 30Hz for 2 min. Samples were spun at 4°C for 15 min at 15000RPM and the supernatant was kept on ice. Bio-Rad DC protein assay kit was used, following the manufacturer's protocol, for protein quantifications. Samples were boiled for 5 min and then subjected to XCell4 SureLock™ Midi-Cell electrophoresis. 20uL of 30ug of protein was run per well. Membranes were blocked for an hour in 5% BSA (in PBS with 0.1% Tween). Primary antibodies were incubated overnight at 4°C, followed by 1hr incubation in horseradish peroxidase–conjugated anti-rabbit secondary antibody diluted in PBS-T 5% milk at room temperature. All images were acquired with Amersham 600 CCD camera and analysed using Image J software. Primary antibodies: S6, S6 phosphorylated (p) S235/S36, and GAPDH.

Rapamycin treatment

We freshly dissolved rapamycin (LC Laboratories) in vehicle solution (5% Tween, 5% PEG, ethanol, distilled water) before use. We administered rapamycin intraperitoneally

once daily at a dose of 5 mg/kg for 3 days before and on the day of collecting tissue for immunoblotting.

Mouse RNA extraction and RNA sequencing

Mouse tissues were homogenised in buffer RLT plus (Qiagen) with β-mercaptoethanol (Sigma, M3148; 10µl/ml) using Qiagen TissueLyser LT, with sterile RNAse-ZAP treated steel beads and operated at 50Hz for 2 minutes. Samples were pre-treated on gDNA eliminator columns and then extracted on RNeasy Plus columns as per manufacturer's protocol (Qiagen), and were immediately snap frozen on dry ice and stored at -80°C. An aliquot of each sample was quantified using 2100 Bioanalyzer (Agilent Technologies). Library preparation was performed by Wellcome Sanger Institute DNA Pipelines. RNA sequencing libraries were prepared using established protocols: library construction (poly(A) pulldown, fragmentation, 1st and 2nd strand synthesis, end prep and ligation) was performed using the NEB Ultra II RNA custom kit (New England Biolabs) on an Agilent Bravo automated system. Samples were sequenced using 75 bp paired-end sequencing reads (reverse stranded) on an Illumina-HTP HiSeq 4000 system. The sequencing data were de-multiplexed into separate CRAM files for each library in a lane. Adapters that had been hard-clipped prior to alignment were reinserted as soft-clipped post alignment, and duplicated fragments were marked in the CRAM files. The data pre-processing, including sequences QC, and STAR alignments were made with a Nextflow pipeline, which is publicly available at https://github.com/wtsi-hgi/nextflow-pipelines/blob/rna_seq_mouse/pipelines/rna_seq.nf, including the specific aligner parameters. We assessed the sequence data quality using FastQC v0.11.8. Reads were aligned to the GRCm38 mouse reference genome (Mus_musculus.GRCm38.dna.primary_assembly.fa, Ensembl GTF annotation v99). We used STAR version 2.7.3a [513] with the --twopassMode Basic parameter. The STAR index was built against Mus_musculus GRCm38 v99 Ensembl GTF using the option -sjdbOverhang 75. We then used featureCounts version 2.0.0 [514] to obtain a readcount matrix. The count data was used as input for differential gene expression analysis using DESeq2 package [515] with SVA correction [516]. The default DESeq2 cut-off of BH-adjusted p-value<0.1 was used for the mouse RNAseq analyses.

For the identification of functionally enriched terms in the differentially expressed genes, Gene ontology (GO) and KEGG pathway enrichment analyses was performed using the grpofiler online suite (http://biit.cs.ut.ee/gprofiler/index.cgi). A threshold of 5% FDR and an enrichment significance threshold of P<0.05 (hypergeometric test with Benjamini-Hochberg False Discovery rate correction for multiple testing) was used. In all analyses, a background comprised of only the expressed genes in the tissue studied (genes where the adjusted p-value yielded a numerical value, different to NA).

A4.1.3 Molecular studies

Molecular studies were performed by Marianna Baybis, Phillip H. Iffland and supervised by Peter Crino.

Mutagenesis and differentiation of KPTN iPSC models

iPSC experiments were all performed in an isogenic background: Kolf2C1_WT (HPSI0114i-kolf_2), feeder-free hiPSC, male, derived from skin tissue using Cytotune 1 reprogramming [517]. In order to create knockout iPSC lines, a guide RNA (gRNA) was selected to target coding exon 8 of the human *KPTN* gene (GRCh38:19:47479882-47479904). The synthetic gRNA along with Cas9 protein were delivered into wildtype Kolf2C1 cells [517] as a pre-complexed ribonucleoprotein via electroporation. The addition of a short single-stranded oligodeoxynucleotide of non-complementary sequence was also added to improve delivery. After a period of recovery, the cells were subcloned and submitted for genotyping by Miseq sequencing. Clones were screened for the presence of frameshift-causing indels and then expanded for banking. Clones identified as unedited in the genotyping analysis were selected as unedited controls, to control for electroporation, selection, and passaging effects. After banking, and subsequent to all differentiation experiments, all cell lines were sent for a second round of confirmatory genotyping. This work was performed by the Gene Editing facility at the Wellcome Sanger Institute.

Differentiation of iPSCs to cortical neural precursor cells

KPTN mutant heterozygous, homozygous edited and wildtype control IPSC lines were cultured in feeder-free conditions as per published protocols [400]. Dual-SMAD inhibition was used in a 10 day neural induction in the presence of XAV to promote regional forebrain identity [400]. After passaging at day 10, cells were replated for

immunocytochemical quality control, and banked at day 14 for RNA extraction and RNAseq analysis. All experiments were confirmed to have high PAX6/GFAP positivity (>90% of cells at day 14).

iPSC RNA extraction and RNA sequencing

RNA extraction and sequencing of iPSC derived materials was performed using manufacturer's protocols for their RNeasy QIAcube kit on a QIAcube automated system (Qiagen). RNA sequencing libraries were prepared using established protocols: library construction (poly(A) pulldown, fragmentation, 1st and 2nd strand synthesis, end prep and ligation) was performed using the NEB Ultra II RNA custom kit (New England Biolabs) on an Agilent Bravo automated system. Indexed multiplexed sequencing was performed on the Novaseq 6000 system (S4 flow cell, Xp workflow; Illumina), collecting approximately 30 million paired-end reads per sample with 100 base read length. The sequencing data were de-multiplexed into separate CRAM files for each library in a lane. Adapters that had been hard-clipped prior to alignment were reinserted as soft-clipped post alignment, and duplicated fragments were marked in the CRAM files. The data pre-processing, including sequences QC, and STAR alignments were made with a custom Nextflow pipeline, which is publicly available at https://github.com/wtsi-hgi/nextflow-pipelines/blob/rna_seq_5607/pipelines/rna_seq.nf, including the specific aligner parameters. We assessed the sequence data quality using FastQC v0.11.8. Reads were aligned to the GRCh38 human reference genome (Ensembl GTF annotation v91). We used STAR [513] version STAR_2.6.1d with the --twopassMode Basic parameter. The STAR index was built against GRCh38 Ensembl GTF v91 using the option -sjdbOverhang 75. We then used featureCounts version 1.6.4 [514] to obtain a readcount matrix. Genes with no count or only a single count across all samples were filtered out. The counts were normalised using DESeq2's median of ratios method [515]. Differential gene expression was analysed using the DESQ2 package [515] with SVA correction [516]. An adjusted p-value threshold of 0.05 was selected to identify significant differences between wildtype and mutant samples.

For the identification of functionally enriched terms in the differentially expressed genes, Gene ontology (GO) and KEGG pathway enrichment analyses was performed using the grpofiler online suite (http://biit.cs.ut.ee/gprofiler/index.cgi). A threshold of

5% FDR and an enrichment significance threshold of P<0.05 (hypergeometric test with Benjamini-Hochberg False Discovery rate correction for multiple testing) was used. In all analyses, a background comprised of only the expressed genes in the tissue studied (genes where the adjusted p-value yielded a numerical value, different to NA).

CMAP Query

To identify anti-correlated transcriptional profiles in the CMAP database [518], we selected the 150 most strongly up- and down-regulated genes from our differential expression analysis of KPTN$^{-/-}$ NPCs, and queried the CMAP via their online tool (https://clue.io/query) using the L1000 TOUCHSTONE (V1.1.1.2) gene expression dataset.

A4.2 Results

A4.2.1 Mouse studies

Generation and characterisation of a Kptn$^{-/-}$ mouse model

These studies used an engineered loss-of-function mouse model generated by the Wellcome Sanger Institute's Mouse Genetics Project [380]. The Kptn$^{tm1a(EUCOMM)Wtsi}$ allele introduces a splice-trapping cassette into an intron downstream of coding exon 8, resulting in an in-frame LACZ open reading frame and premature termination of transcription via an exogenous polyadenylation signal (section A4.1.2, Figure A4.1). The resulting allele closely resembles the truncating allele (p.(Ser259*)) identified in our groups original *KPTN*-related disorder publication [333], resulting in premature termination in Exon 8 and predicted nonsense mediated decay. Homozygous *Kptn$^{-/-}$* mice were observed at expected near-mendelian ratios (see A4.1.2 Mouse studies), and appear grossly normal. Read count quantification in bulk RNAseq from embryonic and adult brain regions shows a robust reduction in *Kptn* mRNA pre- and postnatally, and across adult brain regions (87-90% decrease). This confirms the Kptn$^{tm1a(EUCOMM)Wtsi}$ mouse model is functioning efficiently as a LOF allele, truncating *Kptn* mRNA transcription *in vivo*.

Cognitive deficits in Kptn$^{-/-}$ *mice*

To further characterise the memory deficit identified in the SOR assay (section 4.4.1.2), we tested the *Kptn* mouse model in a spatial memory assay, the Barnes maze [388, 509, 510]. This assay, similar in principle to the Morris water maze, tests a mouse's ability to identify an escape box using spatial cues in an open arena. During both training periods (Training 1 and 2), there was no difference between *Kptn*$^{-/-}$ and wildtype mice in the time taken to approach the escape box (measured by primary latency to approach), indicating that both genotypes were able to locate the target zone with equivalent speed (Figure A4.2A). During the first probe test, after 24 hr (when the escape box was removed), the time spent around the target hole relative to other holes was measured (Figure A4.2B). Both genotypes spent significantly longer around the target hole quadrant when compared to other quadrants ($p \leq 0.05$). However during the second probe test (72 hr after the end of Training 2), the *Kptn*$^{-/-}$ mice do show a deficit in long-term memory: they were on average further from the goal box location (Figure A4.2C, 12.8% further, p=0.0014), and spent significantly less time in the target zone than wildtype controls (Figure A4.2D, 17.9% less time, p=0.0150). These data indicate that *Kptn*$^{-/-}$ mice display deficits in long term hippocampus-dependent spatial memory. The memory impairments identified in *Kptn*$^{-/-}$ mice with SOR and the Barnes maze assays indicate a likely deficit in hippocampus-dependent memory retention.

To test whether the *Kptn*$^{-/-}$ mice also display impairments in tasks less reliant on hippocampal function, we used pairwise discrimination (PD), a visual operant conditioning task using a touchscreen platform [389, 508, 519, 520]. During the PD task, both genotypes selected the correct image approximately 50% of the time in Session 1 (Figure A4.2E, *Kptn*$^{-/-}$:45%; *Kptn*$^{+/+}$:53%), indicating that the mice had no innate bias for either of the images at the start of the task. The average number of days to reach criteria (80% correct trials for two consecutive days) was not significantly different between genotypes, and *Kptn*$^{-/-}$ mice performed no worse than *Kptn*$^{+/+}$ mice (*Kptn*$^{-/-}$: 11.25 days; *Kptn*$^{+/+}$: 14.5 days) (Figure A4.2E,F). We performed a single reversal trial to confirm that the mice were capable of discriminating between images. As expected, the percentage of correct trials dropped significantly below 50% for both genotypes when the CS+ (rewarded) and CS- (non-rewarded) images were swapped (*Kptn*$^{-/-}$:24%; *Kptn*$^{+/+}$:19.2%), indicating a strong association with the original CS+

325

image. These data indicate that *Kptn⁻ᐟ⁻* mice perform at least as well as wildtype animals in the PD task. Taken together, the results of our three independent memory tests suggest that *Kptn⁻ᐟ⁻* mice experience selective deficits, with at least some memory sparing.
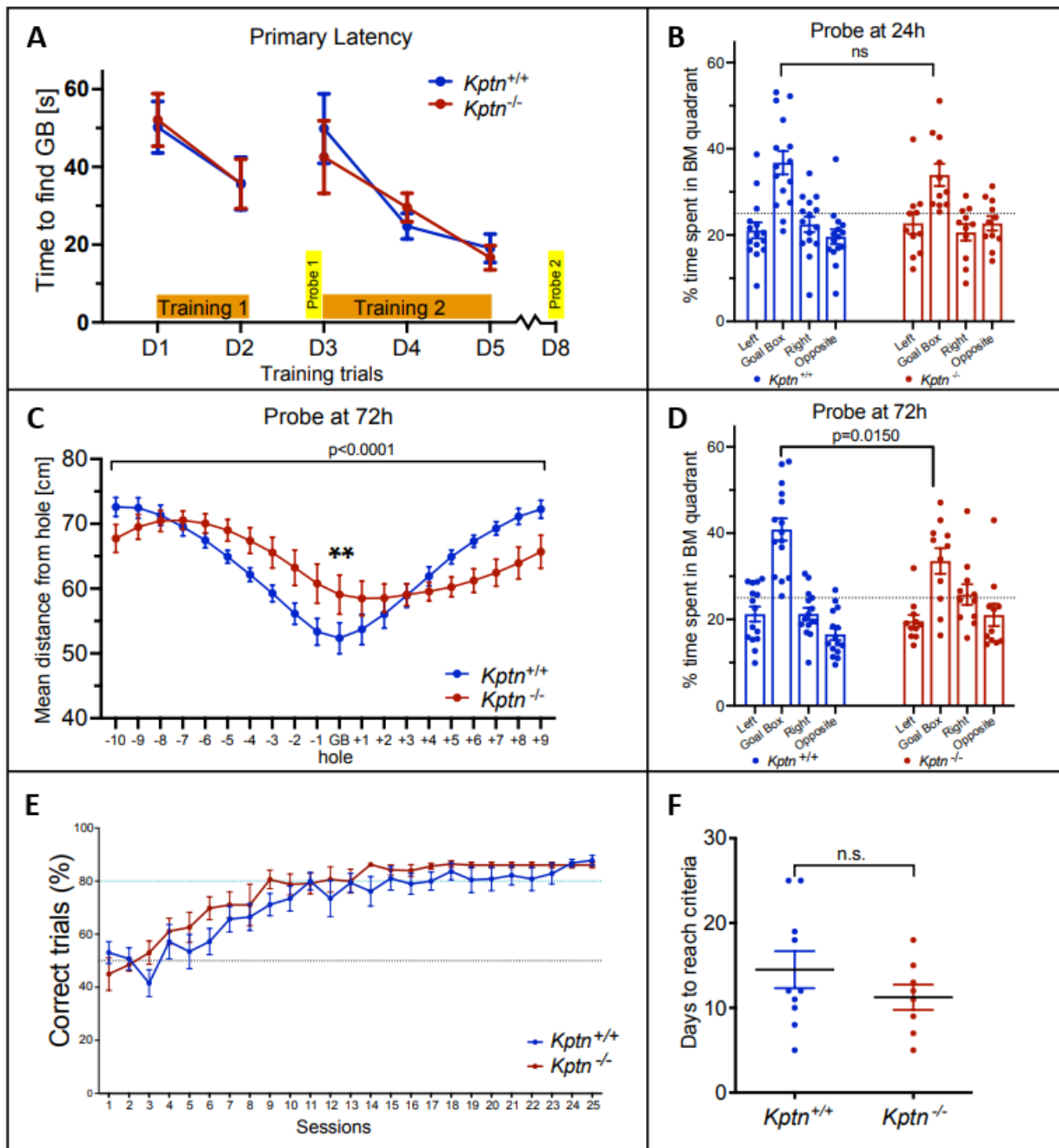
Figure A4.2: *Kptn⁻/⁻* mice have selective memory deficits

(**A-D**) Barnes maze reveals impaired hippocampus-dependent memory acquisition in *Kptn⁻/⁻* mice. (**A**) Primary latency across all days of training (Training 1, D1-D2; Training 2, D3-D5) for *Kptn⁺/⁺* (blue, n=15) and *Kptn⁻/⁻* (red, n=11) mice. There was no genotype difference across all training days and a reduction in time taken to get to the escape box from in Training 1 and Training 2 (2-way ANOVA, interaction GENOTYPE X ZONE $F_{(4,120)}=0.2577$, $p=0.9044$). (**B**) The percentage of time *Kptn⁺/⁺* (blue, n=15) and *Kptn⁻/⁻* (red, n=11) mice spent around the target hole quadrant (Goal Box) and non-target quadrants of the Barnes maze during a probe trial 24hr after Training 1 was similar (2-way ANOVA, interaction GENOTYPE X ZONE $F_{(3,96)}=0.8544$, $p=0.4676$). Both genotypes spent significantly more time at target vs all other holes ($p≤0.05$). (**C**) The mean distance from each hole during probe trial 72hr after

Training 2. $Kptn^{-/-}$ mice were on average further from the target hole compared to $Kptn^{+/+}$ mice (2-way ANOVA, interaction GENOTYPE X HOLE F(19,500)=5.351, p=<0.0001; post-hoc FDR q=0.0045, p=0.0014**). (**D**) The percentage of time $Kptn^{+/+}$ (blue, n=15) and $Kptn^{-/-}$ (red, n=11) mice spent around the target hole quadrant (Goal Box) and non-target quadrants of the Barnes maze during the probe trial 72 hr after Training 2. The $Kptn^{-/-}$ mice spent significantly less time in the target hole quadrant than the $Kptn^{+/+}$ mice (2-way ANOVA, interaction GENOTYPE X ZONE F(3,96)=3.667, p=0.0150; post-hoc FDR q=0.0632, p=0.0150). (**E,F**) $Kptn^{-/-}$ mice do not display an impaired performance in a hippocampus-independent pairwise discrimination task. (**E**) Percentage of correct trials (when CS+ image was nose-poked) out of the total trials completed per session for $Kptn^{-/-}$ (n=8) and $Kptn^{+/+}$ (n=10) mice. Both genotypes started off at around 50% correct trials (black dotted line) in Session 1, which implies the mice had no bias for either image at the start of the task. There was no significant difference between genotypes in learning during the pairwise discrimination (two-way ANOVA, interaction GENOTYPE X SESSION F (1,16)=1.144, p=0.3007). (**F**) There is no significant difference in the number of days to reach criteria between genotypes (p=0.2093, t=1.162 df=16, two-tailed Student's t-test). Values are plotted as mean ± SEM. Data courtesy of Dr Maria Levitin and Dr Sebastian Gerety.

Neuroimaging identifies changes in skull and brain volume in Kptn$^{-/-}$ mice

Microcomputed tomography (µCT) and magnetic resonance imaging (MRI) were used to quantify changes in skull and brain volume [390]. Applying µCT-based cephalometric analysis to examine skull size detected an increase in the height, length, and width of the brain cavity of the skulls of adult *Kptn$^{-/-}$* mice (Figure 4.5A-D). The height of the skull at the midline is consistently elevated throughout the full rostro-caudal extent of the brain cavity (7.7% at Bregma, 4.7% at Lambda, p<0.01), as seen in a sagittal resection of representative 3D volumes (Figure 4.5B). In coronal resections, the height changes appear due primarily to an increase in the dorsal curvature of the frontal and parietal bones (Figure 4.5C,D), rather than an increase in the dorsoventral length of the lateral skull wall. This dorsal bulging is consistent with the changes observed in frontal bossing in humans, where the frontal bone can expand to accommodate the mechanical pressure of a pathological brain overgrowth [391, 392]. Applying MRI tensor-based morphometry followed by voxel-based quantification of brain volume [511], we detected a 7.7%/9.2% increase (female/male, p<0.001) in mean intracranial volume in *Kptn$^{-/-}$* mice at 16 weeks when compared to wild-type controls (Figure 4.5E). Together, these bone and soft-tissue findings indicate that *Kptn$^{-/-}$* mice exhibit both skull shape abnormalities, including overgrowth and deformation, and megalencephaly. These are consistent with the macrocephaly and frontal bossing seen in *KPTN*-related disorder.

Histomorphological analysis of neuroanatomical structures in Kptn$^{-/-}$ mice

Megalencephaly was characterised using previously published histomorphometric methods [324], quantifying neuroanatomical features across the same three coronal brain regions in all subjects. Substantial brain size anomalies were identified in both adult male and female *Kptn$^{-/-}$* mice (16 weeks) when compared to matched wildtypes, with 34 out of 78 parameters affected (Figure A4.3A-C). The total brain area in both sexes was significantly increased across the three coronal sections (+7.9-14.4%, p≤0.0007), concomitant with enlarged cortices in all areas measured, and enlarged corpus callosum structures (Figure 4.5F,G). The lateral ventricles were the only brain regions exhibiting a decreased size (section 1: -36%, p=0.05; section 2: -54%, p=0.02), likely indicating a compensatory reaction to increased brain volume constrained within the skull. Interestingly, no significant difference in hippocampus-

related size parameters was detected (Figures 4.5G and A4.3A-C). The histological findings are consistent between sexes (Figure A4.3A-C). Of note, brain enlargement is accompanied by increased cell number, rather than a change in cell body size (Table A4.1) as has been observed in some models of mTOR related macrocephaly [521]. Taken together, the inactivation of Kptn results in skull shape changes, and megalencephaly, with overgrowth affecting most cortical regions. Both the enlargement and deformation of the skull and brain are consistent with those observed in human patients with *KPTN*-related disorder.

| Barcode	All_L	All_cellcount_L	All_cellarea_L	All_cellcirc _L	All_cellsol _L
2_M01819898	0.018838715	3885	0.004906475	0.774949194	0.90167562
2_M01819915	0.022363431	4101	0.004974723	0.782434199	0.904654797
2_M01819917	0.0021358	4062	0.005255243	0.79895173	0.912727664
2_M01819936	0.023110032	4008	0.005661338	0.777894988	0.9054395
2_M01820006	0.022858053	3755	0.004846617	0.803658852	0.913518381
2_M01820065	0.019576468	3346	0.004088987	0.806899908	0.913101299
2_M01819785	0.025429793	3946	0.005312496	0.787148916	0.908153227
2_M01819914	0.025529888	NM	NM	NM	NM
2_M01819997	0.026982741	4415	0.005774248	0.785429924	0.908782053
2_M01820022	0.02618036	4242	0.005675851	0.762144705	0.89613312
2_M01820023	0.025743827	4063	0.005429067	0.791742	0.9107689
2_M01820025	0.026375081	4366	0.005815022	0.772234599	0.90181333
T-TEST	0.037488524	0.046506685	0.030105043	0.20089249	0.342775805
AVG_WT	0.018147083	3859.5	0.004955564	0.790798145	0.908519544
AVG_HOM	0.026040282	4206.4	0.005601337	0.779740029	0.905130126
%change	143.495687	108.9882109	113.0312713	98.6016512	99.6269296

Table A4.1: Quantification of neuroanatomical cellular parameters in *Kptn⁻ᐟ⁻* mice

Raw measurements for adult male mice at 16 weeks of cellular parameters including cell count (cellcount), total cell area (cellarea), cell circularity (cellcirc) and cell solidity (cell sol). p-values are from 2-sample two-tailed Student's t-tests.
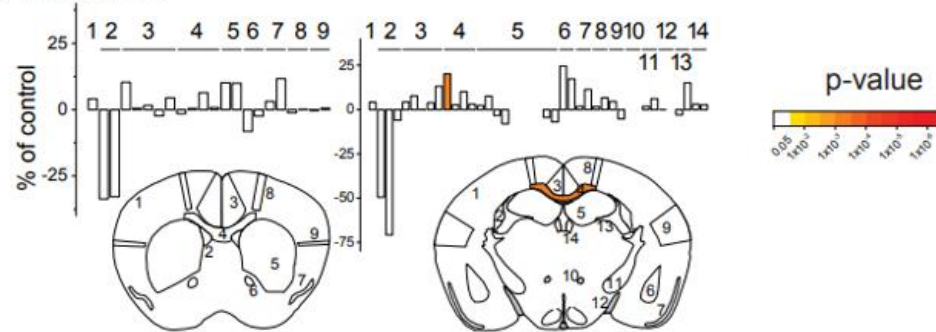
A Males 16wks
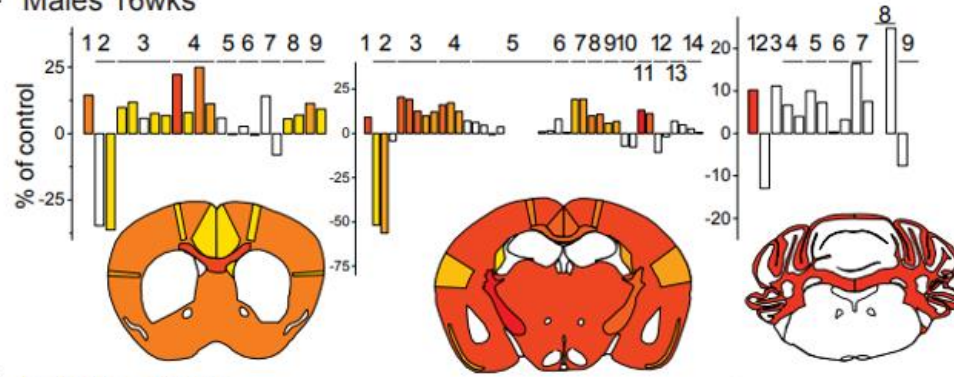
B Females 16wks
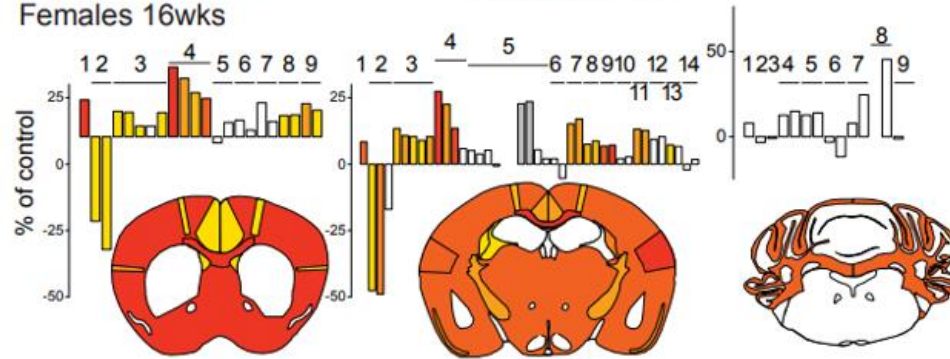
D Males P20

C

Section1

1 Total Brain Area
2 Lateral Ventricle_Left side
   Lateral Ventricle_Right side
   Cingulate cortex area_Left side
   Cingulate cortex area_Right side
3 Cingulate cortex area_Width_Left side
   Cingulate cortex_Width_Right side
   Cingulate cortex_Height
   genu of the corpus callosum
4 genu of the corpus callosum_Height
   genu of the corpus callosum_Width_Basal
   genu of the corpus callosum_Width_Top
5 caudate putamen_Left side
   caudate putamen_Right side
6 anterior_commissure_Left side
   anterior_commissure_Right side
7 piriform cortex_Left side
   piriform cortex_Right side
8 primary motor cortex_Left side length
   primary motor cortex_Reft side lengthcc
9 secondary somatosensory cortex_Left side length
   secondary somatosensory cortex_Right side length

Section 3

1 Total Brain Area
2 fourth Ventricle
3 pons
4 pyramidal_L
   pyramidal_R
5 genu facial nerve_L
   genu facial nerve_R
6 cochlear nucleus_L
   cochlear nucleus_R
7 lateral dentate Cb_L
   lateral dentate Cb_R
8 interposed cerebellar_L
   interposed cerebellar_R
9 internal granule layer
   folia number

Section 2

1 Total Brain Area
   Lateral Ventricle_Left side
2 Lateral Ventricle_Right side
   dorsal 3rd ventricle
   retrosplenial granular cortex_Left side
   retrosplenial granular cortex_Right side
3 retrosplenial granular cortex_Width_Left side
   retrosplenial granular cortex_Width_Right side
   retrosplenial granular cortex_Height
   corpus callosum
4 corpus callosum_Height
   ccorpus callosum_Width
   dorsal hippocampal commissure
   hippocampus
   pyramidal layer
   dentate gyrus_Left side
   dentate gyrus_Right side
   lacunosum moleculare_Left side
5 lacunosum moleculare_Right side
   radiatum layer_Left side
   radiatum layer_Right side
   oriens layer_Left side
   oriens layer_R
6 amygdala_Left side
   amygdala_Right side
7 piriform cortex_Left side
   piriform cortex_Right side
8 primary motor cortex_Left side length
   primary motor cortex_Right side length
9 secondary somatosensory cortex_Left side length
   secondary somatosensory cortex_Right side lengtl
10 mammillothalamic tract_Left side
   mammillothalamic tract_Right side

11 internal capsule_Left side
   internal capsule_Right side
12 optic tract_Left side
   optic tract_Right side
13 fimbria_Left side
   fimbria_Right side
14 med habenular_Left side
   med habenular_Right side

p-value

E

□WT ■*Kptn-/-*
* : pval < 0.05
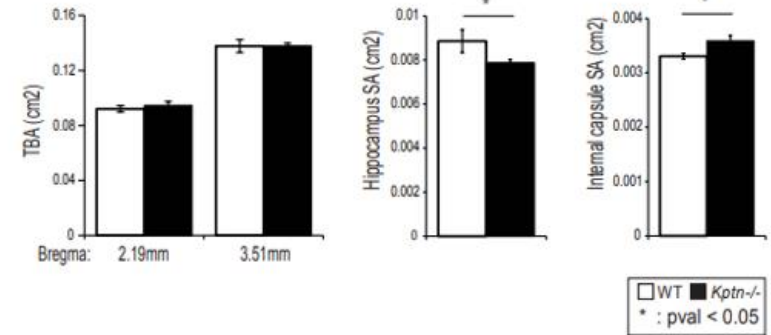
Figure A4.3: Inactivation of *Kptn* gives rise to major neuroanatomical defects

(**A**) Neuroanatomical findings in adult male *Kptn* mice (n=8 *Kptn*$^{-/-}$ versus 8 *Kptn*$^{+/+}$ controls). Below, a representative heatmap of the p-values for the three studied sections (Bregma +0.98mm, Bregma -1.34mm and Bregma -5.80mm). Above, histograms of the percentage change of *Kptn*$^{-/-}$ relative to *Kptn*$^{+/+}$ controls (100%). (**B**) Neuroanatomical findings in adult female *Kptn* mice (n=8 *Kptn*$^{-/-}$ versus 8 *Kptn*$^{+/+}$ controls), as described for males in (**A**). In the legends in (**C**), green text indicates length measurements and black text denotes area measurements. (**D**) Neuroanatomical findings in three week old (P20) male *Kptn* mice (n=6 *Kptn*$^{-/-}$ versus 6 *Kptn*$^{+/+}$ controls), as described for adults in a, with legend in (**C**). (**E**) At Birth (P0), the total brain area (tba) of *Kptn*$^{-/-}$ mice (n=9) is unchanged (2.77% increase, p=0.494) compared to *Kptn*$^{+/+}$ controls (n=8). Only the hippocampus (11% increase, p=0.0489) and internal capsule (8.6% decrease, p=0.0361) show statistically significant changes at this stage of development. All p-values are from 2-sample two-tailed Student's t-tests. Data courtesy of Dr Sebastian Gerety and Dr Binnaz Yelcin.

To better understand the developmental trajectory of brain overgrowth in the mouse model, histomorphology was compared at birth (Postnatal day P0), P20, and 16 weeks of age. P0 allows an assessment of whether *Kptn*-/- mice are born with enlarged brains, or conversely overgrow postnatally. P20 represents the stage at which wildtype mice reach 90-95% of adult brain weight [522, 523] and thus allows us to distinguish between an increased growth rate and an extended brain growth period. In the former, we would anticipate a larger brain by the end of normal brain growth (P20), while the latter would lead to later appearance of overgrowth (16 weeks). Using histological methods described above, no phenotypic differences at P0 were detected in the vast majority of parameters (51 out of 53) including the total brain area (Figure A4.3E), indicating that the megalencephaly phenotype is associated with postnatal developmental processes. Interestingly, two parameters were altered at P0: the total area of the hippocampus is reduced by 11% in the *Kptn*-/- mice (p=0.04), and the internal capsule increased by 9% (p=0.03), suggesting that some morphological anomalies originate from prenatal stages but these are highly tissue restricted (Figure A4.3E).

A similar analysis on P20 mice identifies very few statistically significant changes in regional brain size (Figure A4.3D), despite a broad upward trend. Consistent with all stages examined, the ventricles of P20 *Kptn*-/- mice are reduced in size, supporting the idea that there is ongoing, but mild overgrowth at early stages that continues into adulthood, eventually resulting in severe megalencephaly. It is striking that *Kptn*-/- mice lack significant overgrowth at P20, when wildtype mice will have reached the end of normal brain growth. The progressive nature of the disorder is well illustrated when we plot measurements of total brain area, motor cortex, and hippocampus with time (Figure 4.6). The underdevelopment of the hippocampus at P0 recovers with age, but never reaches statistically significant overgrowth during the life of the mouse. Taken together with the significant brain size changes in adult *Kptn*-/- mice, these data indicate that the megalencephaly seen in the *Kptn* model is postnatal, progressive, and results from an extended period of brain growth beyond the normal period.

A4.2.2 Transcriptomic studies

Transcriptional changes in Kptn-/- *mouse and human induced pluripotent stem cell (iPSC) models highlight mTOR pathway dysregulation*

While few transcriptional readouts have been described for mTOR activity, the tight regulation of this critical metabolic pathway involves extensive use of feedback loops to maintain homeostasis [524]. We hypothesised that the loss of negative mTOR regulation in our *Kptn*-/- model would likely result in detectable changes to the expression of these mTOR pathway members, and genes involved in the downstream cellular processes. We performed RNA-seq analysis on P21 and adult brain tissue, comparing gene expression between *Kptn*-/- and *Kptn*+/+ animals. We recorded extensive dysregulation of both positive and negative regulators of the mTOR pathway (Figure A4.4A). Interestingly, the hippocampus and cortex show the strongest dysregulation, concordant with the immunostaining results in adult tissue. A notable observation is the enrichment in overexpression of negative mTOR regulators (Figure A4.4A): this likely reflects a strong negative feedback response attempting to attenuate the hyperactivated mTOR signalling resulting from loss of Kptn protein. Numerous genes involved in protein synthesis, including most ribosomal protein genes, are dysregulated in the *Kptn*-/- mice (Figure A4.4B). These two observations are supported by significant enrichment of terms related to these processes, when assessed by Gene Ontology enrichment [525]. When tested specifically for dysregulation of mTOR pathway genes, 5 out of 6 tissues examined show a significant enrichment (hypergeometric test $p < 0.05$). These data are concordant with the known regulation of ribosomal protein genes by mTOR and its critical role in controlling protein synthesis [526-528].
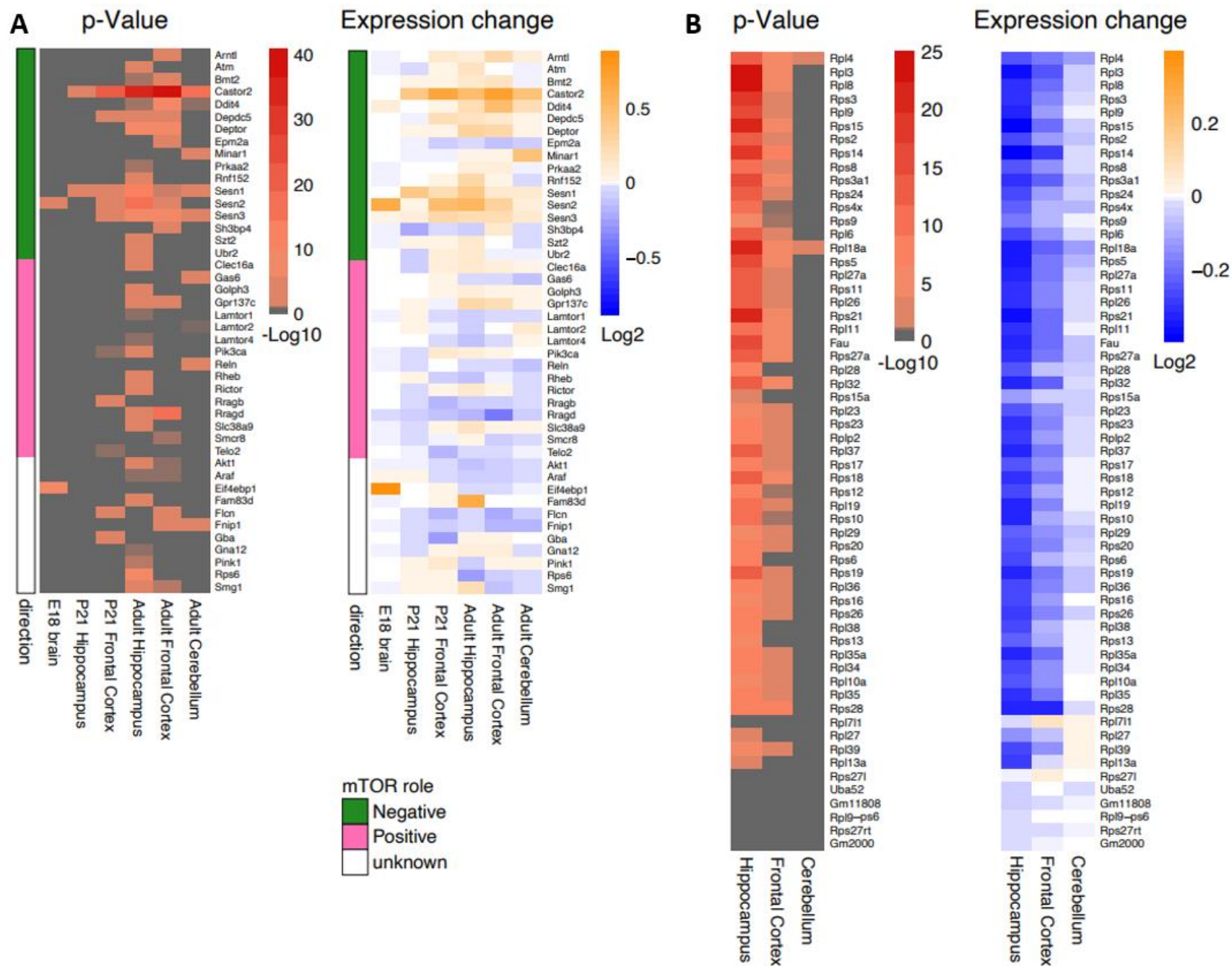
Figure A4.4: Kptn loss-of-function results in dysregulated expression of mTOR pathway genes

(**A,B**) transcriptomic changes to mTOR pathway components (**A**) and downstream ribosomal gene network (**B**) in *Kptn*$^{-/-}$ mice at embryonic and postnatal stages as indicated. Statistically significant ($\alpha<0.05$) LOG2-fold changes in expression are indicated by non-grey adjusted p-value heatmap cells. Role in mTOR pathway is indicated as green/pink/white colour scheme in (**A**). Data courtesy of Dr Maria Levitin and Dr Sebastian Gerety.

To confirm the relevance of our mouse findings to human biology, human iPSC models of *KPTN*-related disorder were created and differentiated into PAX6-positive cortical neural precursor cells (NPCs) [400] as described in A4.3 (Figure A4.5A). Transcriptomic analysis of these cortical NPCs and isogenic wildtype controls reveals significant differential expression of numerous genes involved in mTOR signalling, and nearly complete dysregulation of the ribosomal protein gene network (Figure A4.5B,C). The TSC proteins, important inhibitors of the mTOR pathway, are both downregulated in *KPTN*$^{-/-}$ NPCs. LOF variants in TSCs are associated with tuberous sclerosis complex, a multi-system tissue overgrowth syndrome affecting brain, skin, kidney, heart, and lung [529]. 36% of genes dysregulated in *TSC2*$^{-/-}$ neurons [530] are also differentially expressed in *KPTN*$^{-/-}$ NPCs (301 out of 834 expressed in NPCs, 1.35 fold over-enriched, p=4.3x10^{-10}), thus identifying a common signal in neural cell types driven by convergent disease mechanisms affecting the mTOR pathway.

The Connectivity Map (CMAP [518]) is a large-scale collection of transcriptional response profiles after compound exposure (20,000 different compounds and drugs) or gene knockdown across numerous cell lines. The stated aim of the project is to identify potential therapeutics to reverse disease states, through the identification of drug-response profiles that are negatively correlated with disease profiles. The CMAP database was queried with the transcriptional response signature of the *KPTN*$^{-/-}$ NPCs. Torin-2 [531], a known mTOR inhibitor, scores as the most strongly negatively correlated profile (global correlation score of -96.28 across all cell lines), and is therefore considered the most promising treatment for the queried disease-profile. Interestingly, in HCC515 cells five of the top 32 anti-correlated signatures are mTOR inhibitors (torin-2, QL-X-138, torin-1, sirolimus, and WYE-354) with scores between -99.56 and -93.23. This strong anticorrelation with mTOR inhibitor responses, detected in non-neural cancer lines, suggests that the effect of loss of KPTN protein on the mTOR pathway is robust and specific.

Figure A4.5: Human neural stem cell model of *KPTN*-related disorder has mTOR activation and developmental disorder gene dysregulation
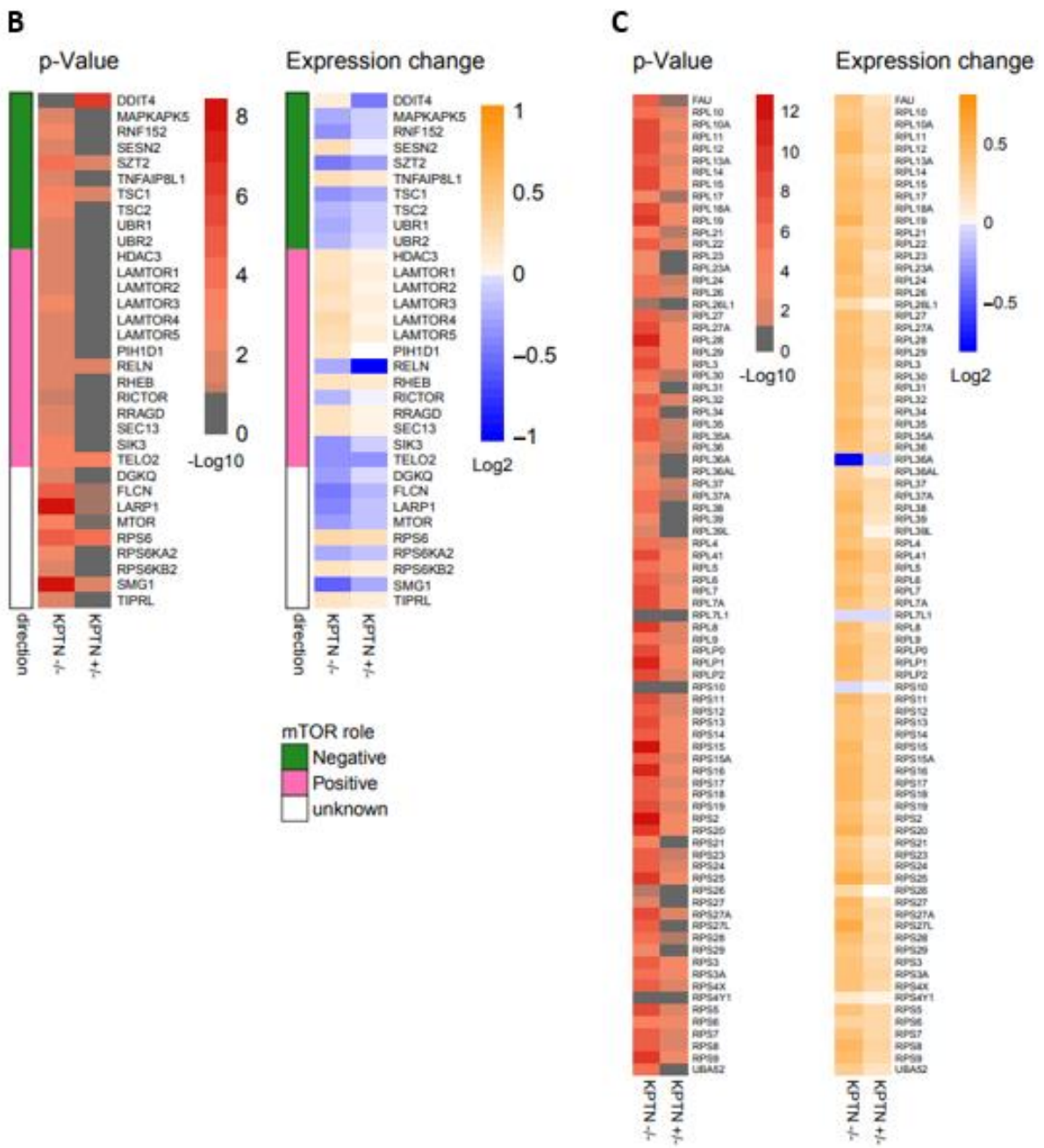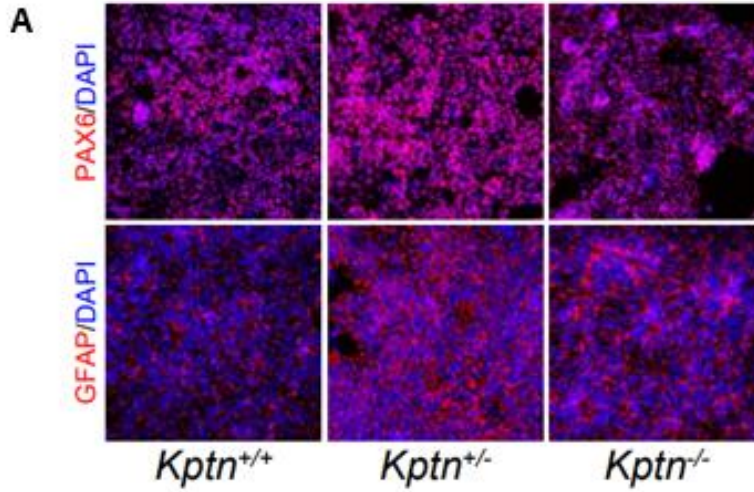
(**A**) Engineered KPTN mutant IPSCs differentiated into cortical neural stem cells, stained for PAX6 (top) and GFAP (bottom) neural stem cell markers in red. (**B,C**) KPTN LoF NPC models display changes in gene expression of mTOR pathway components (**B**) and downstream ribosomal gene network (**C**). Heterozygous KPTN mutant cells (*KPTN*+/-) have an intermediate mTOR pathway phenotype consistent with dosage sensitivity to loss of KPTN expression. (**D**) Mouse brain and (**E**) human neural cell models of KRD show significant dysregulation of dominant haploinsufficient developmental disorder associated genes (statistically significant LOG2-fold changes in expression are indicated by non-grey adjusted p-value heatmap cells). Data courtesy of Dr Maria Levitin and Dr Sebastian Gerety.

Seizure gene networks in KPTN-related disorder

While not directly tested in our mouse model, we did not observe any clinical seizures during the maintenance and testing of *Kptn*$^{-/-}$ mice. The absence of detectable spontaneous seizures in this mouse model is not totally unexpected: Other mouse models of human seizure genes such as *SZT2* (a KPTN binding partner and mTOR regulator) [367] show a reduced threshold for seizures after pentylenetetrazole and electroconvulsive treatment, while lacking the spontaneous seizures observed in the corresponding human disorder [355, 356]. Transcriptomic analysis of *KPTN*-related disorder models, however, reveals the dysregulation of numerous known epilepsy genes [532, 533] (syndromic and non-syndromic) in both mouse brain (290 out of 844 expressed) and cortical NPCs (281 out of 807 expressed, 1.3 fold over-enriched p=8.8x10^{-8}). Among the differentially expressed seizure genes, around half of the mouse genes (n= 122) and a third of the human genes (n= 97) are concordant between direction of change and known mechanism of disease (e.g. downregulated expression of a LOF mechanism disease-associated gene). There is significant overlap in differentially expressed seizure genes between mouse and human *KPTN*-related disorder models (107 genes in common, 1.5 fold over-enriched, p=3.6x10^{-7}). Of particular interest are those known to be dosage sensitive, and are therefore most likely to contribute to seizure predisposition. For example, *SCN1A*, *SCN8A*, *SLC2A1*, *SYNGAP1*, *SHANK1*, and *LGI1* all show more than 30% reduction in expression in *KPTN*$^{-/-}$ NPCs. KCNQ3, a potassium channel protein found mutated in neonatal seizures, shows a 25% increase in expression in these cells, consistent with the suspected altered-function mechanism in patients [534, 535]. Of others known to operate by an altered or gain-of-function mechanism, *Sik1* is of note, with a 90% increased expression in the frontal cortex of *Kptn*$^{-/-}$ mice (p-adj = 0.0012). Patients carrying *SIK1* missense mutations experience severe developmental epilepsy, with mutations altering the protein's ability to correctly regulate MEF2C activity in neurons [536, 537]. Taken together, these data suggest that KPTN LOF is resulting in dysregulation of numerous genes that control electrophysiological homeostasis in brain tissue, leading to a susceptibility to seizures in affected individuals.

Increases in radial glial and intermediate precursor marker expression

The mTOR pathway is both active in, and critical for the development and differentiation of neural stem cells in both mice and humans [360, 361, 538]. Pathological increases in mTOR signalling, seen in TSC and somatic mTOR mutations inducing hemimegancephaly [364, 366] point to a significant input of the mTOR pathway in the control of neural stem cell expansion and differentiation. To understand how the loss of KPTN protein is leading to hypercellularity in *KPTN*-related disorder, expression of markers of neural stem cells, radial glia and proliferating neuronal progenitors were examined in our mouse model. Significant increases in levels of radial glial markers (e.g. Nestin, Vimentin, GLAST/Slc3a1, and Thrsp) and intermediate progenitor markers (e.g. Tbr1/Eomes, Neurod2, Btg2) [539-543] were identified in the mouse brain at both P21 and adult stages. These striking increases in gene expression support a model in which the persistence, beyond P21, of excessive neural stem cells and cycling neural progenitors contribute supernumerary neuronal cells, resulting in significant pathological brain overgrowth. These data provide an important starting point for further work on the cellular basis of megalencephaly in *KPTN*-related disorder.

Dosage sensitive gene networks are disrupted in KPTN-related disorder models

The regulation of mTOR signalling is critical to properly regulated growth and differentiation during development [352, 544]. The loss of an important negative regulator of this pathway during development would therefore be expected to disrupt important gene networks that regulate these processes. Indeed, numerous genes causally linked to other neurodevelopmental disorders are differentially expressed in the *KPTN* models. Haploinsufficient genes (dosage sensitive disease associated genes, Figure A4.5D,E), whose phenotype appears in heterozygous loss-of-function, are key regulators of brain development and normal brain function [545] and there is consistent downregulation (82% of 111 genes, Figure A4.5E) of these genes in *KPTN*-/- human cortical neural precursors (p<0.00001, Fisher's exact test). This includes a more than 25% reduction in expression of important chromatin modifying proteins and transcriptional regulators AUTS2, AHDC1, ANKRD11, CHD7, CREBBP, CHD2, EBF3, KMT2D, TBR1, TCF20, SETD1B and ZIC2. Individual disruptive mutations in any one of these genes results in severe neurodevelopmental deficits in *KPTN*-related disorder associated phenotypic areas. Substantial changes in the expression of such important proteins indicates a significant impact of *KPTN* gene function on a broad network of

developmentally important genes, whose reduced expression has a known impact on cognitive function.

5 APPENDIX: Clinical features and additional genomic variants identified in affected individuals with *INPP4A*-related neurodevelopmental disorder

A5.1 Clinical features of affected individuals with *INPP4A*-related neurodevelopmental disorder

Family 1 comprises a Pakistani family with two affected female siblings (IV:4, IV:5, Figure 5.2A, Family 1) aged 25 and 21 years. Both sisters were born following a normal full-term pregnancy with no complications and presented within the first year of life with lower limb paralysis and global developmental delay. Motor development is significantly impaired and neither individual is able to walk; Individual IV:4 is mobile by crawling using hands and feet, whereas Individual IV:5 has a more severe phenotype and is unable to crawl or sit without support. Both have significant expressive and receptive speech impairment, Individual IV:5 is only able to make sounds and has very limited understanding, whereas individual IV:4 is able to say 10-15 words and has some understanding. Although neither has had a formal IQ assessment, both have a severe degree of intellectual disability. Reduced growth parameters are apparent with significant short stature and microcephaly (Table 5.2). Both sisters are hyperteloric and have bilateral strabismus (Figure 5.3A-D). Individual IV:4 has dental anomalies including oligodontia with absent central incisors. Behavioural abnormalities include sleep disturbance, reduced social interaction and repetitive head shaking.

Family 2 entails six affected siblings from an Iranian family (Figure 5.2A), who are related to and live in the same village as the affected individuals from the Iranian family previously published by Banihashemi *et al* [466] (reviewed below). Limited clinical details are available for three deceased affected siblings, although they are reported to have had similar features to their living affected siblings, and their cause of death is unknown. Three male siblings (aged 34-40) have severe intellectual disability (IQ 35-45) with global developmental delay, are non-ambulatory (all sit unsupported, one brother can stand with support) with wasting and weakness of the lower limbs, severe speech impairment and limited understanding (Table 5.2). Other more variable features include oligodontia with absent central incisors, raised BMI, happy facial expression and pitting of the skin over the posterior scalp (2/3). Reassessment of the

six affected Iranian individuals originally described by Banihashemi *et al* [466] identified several previously unreported features, including repetitive head shaking (2/6) and syndactyly (1/6) (Table 5.2).

Families 3 and 4 each comprised of a single male Palestinian individual, aged 27 and 7.5 years respectively (Figure 5.2A) are both from the same village in Jerusalem, although are not known to be closely related. Both have a similar clinical phenotype including microcephaly, profound intellectual disability and developmental delay (non-ambulatory and never achieved sitting, non-verbal with no communication) poor feeding from birth requiring gastrostomy, drooling and have severe intractable tonic and GTC seizures with pharmocoresistance (Table 5.2). Both individuals have severe spastic quadriparesis with opisthotonus and features of movement disorder including choreoathetosis and dystonic movements. Other features include joint contractures and poor visual acuity and Individual II:1 from Family 3 also has a tracheostomy. Facial features are described as coarse with prominent ear lobes. Neuroimaging (CT in Individual II:1, Family 3 and MRI in Individual II:1, Family 4) abnormalities of both individuals revealed cerebellar hypoplasia, cerebral atrophy and mega cisterna magna, with Dandy-Walker malformation and hyperdense areas in bilateral parietooccipital cortex and subcortical white matter also reported in Individual II:1 (Family 3). An EEG of Individual II:1 from Family 4 showed high voltage spike and wave abnormality in the right parietal region; and effective antiepileptic medication includes valproate, topiramate, leviteracitum and clonazepam for these individuals.

Family 5 entails a 13 year old Egyptian female singleton (Figure 5.2A) with short stature, microcephaly (-4.26SD), severe to profound intellectual disability, spastic quadriparesis with acquired arthrogryposis (wrists, elbows, knees), agitation, poor sleep and encephalopathy with severe intractable seizures of GTC, myoclonic and tonic type, with developmental regression at 1 year. Cerebellar features include nystagmus and intention tremor. Multiple antiepileptic medications including lamotrigine, leviteracitum, clonazepam, valproate, clonazepam and topiramate. EEG shows bitemporal epileptogenic discharges. MRI brain imaging shows a large temporo-parietal cyst, progressive vermian atrophy, mild cortical atrophy with wide inter-hemispheric fissure and mildly dilated asymmetric lateral ventricles, thin corpus callosum and mild thinning of the brainstem (Figure 5.3G-H). Facial features include

a long face with high forehead and anterior hairline, prominent nose, short philtrum, open mouth, broad chin and low set ears.

Two siblings (male Individual IV:1, and female Individual IV:2, aged 10 and 7 years respectively) from an interrelated Iranian family comprise Family 6 (Figure 5.2A) who have features of moderate to severe intellectual disability, global developmental delay, cerebellar features of ataxia, tremor and nystagmus, and autistic features. Individual IV:1 walked age 7, and currently is able to run, although has difficulty climbing the stairs and speak only two words. Individual IV:2 is non-ambulatory and is able to crawl only and speaks only three words. MRI neuroimaging revealed cerebellar atrophy and thin corpus callosum in both siblings (Figure 5.3E,F).

Family 7 comprises a 21 year old affected Algerian male (Figure 5.2A) with severe intellectual disability and speech impairment (speaks a few words only), mild delay in motor development (walked at 2 years) and epilepsy (generalised tonic-clonic and absence seizures) with onset at 11 years and pharmacoresistance, currently controlled with lamotrigine, sodium valproate and clobazam. Mild cerebellar features are present with intention tremor (onset 2-3 years) and wide-based gait, although no strabismus was reported. Behavioural features include agitation, aggressive episodes and social withdrawal. This individual has macrocephaly (OFC 62.5cm, +3.06 SDS) with frontal bossing and a large jaw, although the father also has reported macrocephaly with no other features. MRI brain performed at 15 years identified vermian atrophy.

Family 8 entails a 6-year-old singleton female from Pakistan (Figure 5.2A) with moderate intellectual disability, severe speech impairment (speaks only "mama" and "papa" or sounds), mild motor delay (walked at 2 years), left-sided hemiparesis and generalised tonic-clonic seizures from the age of 2.5 years controlled with sodium valproate. Mild cerebellar features include strabismus and unsteady gait, with behavioural features of impaired social interaction and aggression. Subtle dysmorphic features include macroglossia, downslanting palpebral fissures, midface hypoplasia and joint hypermobility. Neuroimaging results are currently awaited.
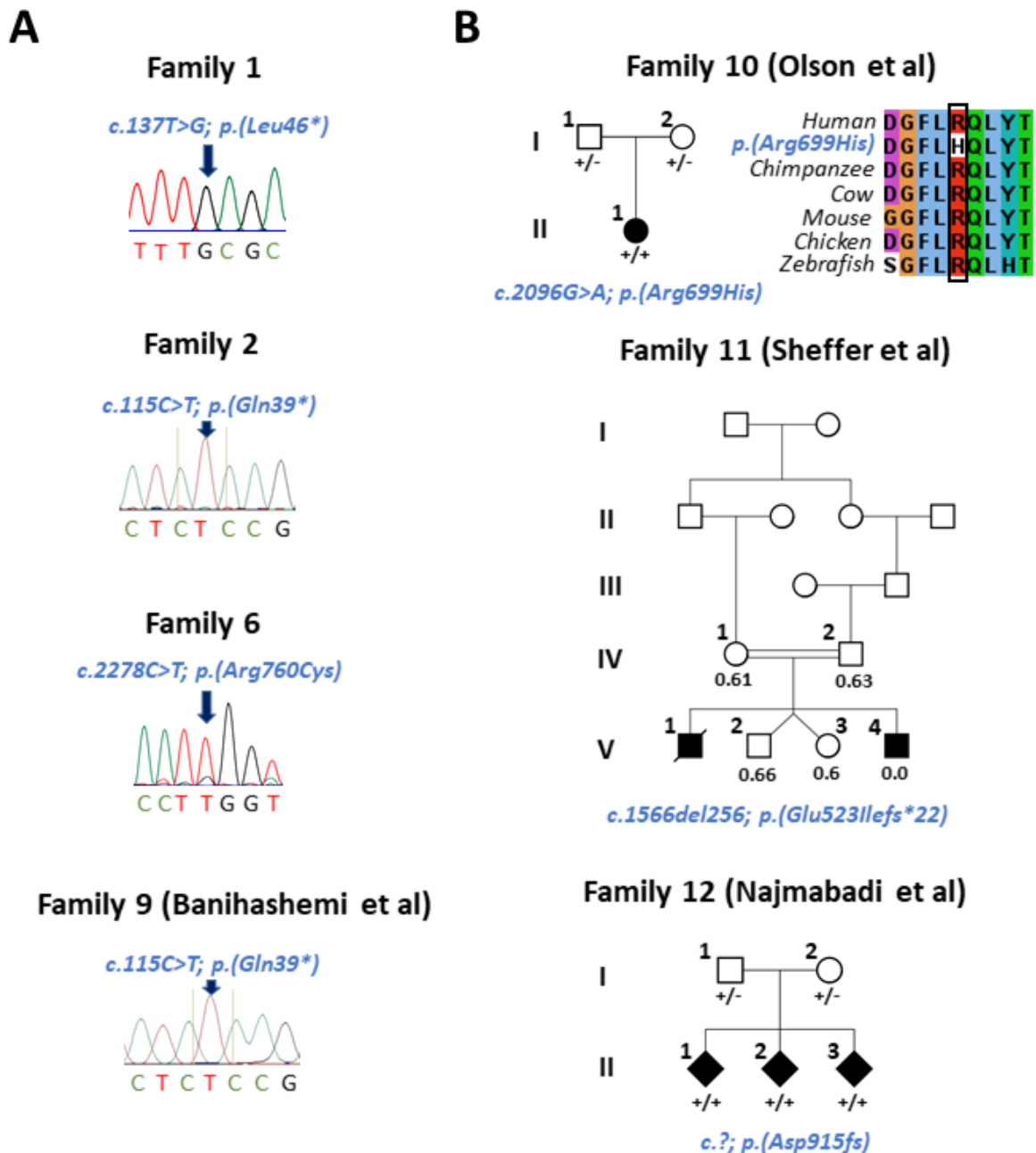
Figure A5.1: Dideoxy sequencing chromatograms of *INPP4A* variants and pedigrees of previously published families

(**A**) Dideoxy sequencing chromatograms for biallelic *INPP4A* variants identified in families 1, 2, 6 and 9. (**B**) Family pedigrees showing segregation of *INPP4A* variants for previously published individuals with *INPP4A*-related disorder (Families 10-12). Family 10 published by Olson *et al* includes a multiple species amino acid alignment for the c.2096G>A; p.(Arg699His) variant showing evolutionary conservation to Zebrafish [465]. The pedigree for Family 11 (Sheffer *et al*) shows segregation of a homozygous 256bp deletion, with numbers in the pedigree indicating the relative dosage (0.0 indicates homozygosity, 0.6 heterozygosity) [464]. Family 12 (Najmabadi *et al*) show segregation of a frameshift variant, although full variant details were incomplete [463].

Family	Gene	Zygosity	GRChr37:g	c.Nomen	p.Nomen	gnomAD v2.1.1 MAF	gnomAD v3.1.2 MAF	In silico predictions		
								Polyphen	SIFT	REVEL
1	TPR	Hom	Chr1:g.186337104G>A	NM_003292.2: c.341C>T	NM_003292.2: p.(Thr114Ile)	0.00008137	0.00003944	Benign	Deleterious	0.084
	LMF2	Hom	Chr22:g.50943509C>T	NM_033200.2: c.1231G>A	NM_033200.2: p.(Val411Met)	0.0002116	0.00008540	Possibly damaging	Deleterious	0.259
	SAV1	Hom	Chr14:g.51101927C>T	NM_021818.3: c.1126G>A	NM_021818.3: p.(Ala376Thr)	0.0002230	0.00001972	Benign	Tolerated	0.162
	ULK4	C Het*	Chr3:g.41831305T>C	NM_017886.3: c.2041A>G	NM_017886.3: p.(Thr681Ala)	0.00002020	0.000006575	Benign	Tolerated	0.046
	ULK4	C Het*	Chr3(GRCh37): g.41961277C>T	NM_017886.3: c.575G>A	NM_017886.3: p.(Arg192Lys)	-	0.00001314	Benign	Tolerated	0.071
	SPSB2	C Het*	Chr12:g.6981879C>T	NM_032641.3: c.187G>A	NM_032641.3: p.(Gly63Arg)	0.00001202	0.000006567	Possibly damaging	Tolerated	0.139
	SPSB2	C Het*	Chr12:g.6982059G>A	NM_032641.3: c.7C>T	NM_032641.3: p.(Gln3*)	0.00002771	0.00002629	N/A	N/A	N/A
	FRAS1	C Het*	Chr4:g.79204114C>A	NM_025074.7: c.1248C>A	NM_025074.7: p.(Cys416*)	-	-	N/A	N/A	N/A
	FRAS1	C Het*	Chr4:g.79239999G>A	NM_025074.7: c.1996G>A	NM_025074.7: p.(Ala666Thr)	0.00007683	0.00001972	Benign	Tolerated	0.042
	FRAS1	C Het*	Chr4:g.79240083G>A	NM_025074.7: c.2080G>A	NM_025074.7: p.(Glu694Lys)	0.00006823	0.00005258	Benign	Tolerated	0.057
	PKHD1L1	C Het*	Chr8:g.110418599C>A	NM_177531.5: c.1705C>A	NM_177531.5: p.(Gln569Lys)	0.00004773	-	Benign	Deleterious	0.173
	PKHD1L1	C Het*	Chr8:g.110539235T>C	NM_177531.5: c.12707T>C	NM_177531.5: p.(Leu4236Ser)	0.00006944	0.00003285	Benign	Tolerated	0.45
	ADAMTS7	C Het*	Chr15:g.79058319G>C	NM_014272.4: c.3934C>G	NM_014272.4: p.(Pro1312Ala)	0.00009731	0.00007969	Benign	Tolerated	0.12
	ADAMTS7	C Het*	Chr15:g.79068673C>A	NM_014272.4: c.1563G>T	NM_014272.4: p.(Trp521Cys)	0.000008034	-	Probably damaging	Deleterious	0.606

5	*SLC43A2*	Het	Chr17:g.1494606_1494622dup	NM_001284498.1: c.872_888dup	NM_001284498.1: p.(Ser297Alafs*13)	-	-	N/A	N/A	N/A
	DOT1L	Het	Chr19:g.2222242A>T	NM_032482.2: c.3074A>T	NM_032482.2: p.(Lys1025Met)	0.000004031	-	Probably damaging	Deleterious	0.067

Table A5.1: Genomic variants identified by WES in each family that could not be fully excluded

Abbreviations: MAF, mean allele frequency; Hom, homozygous; Het, heterozygous; C Het, compound heterozygous; N/A, not applicable.

6 APPENDIX: Publications relating to the work in these PhD studies

Rawlins LE*, Almousa H*, Khan S*, Collins SC*, Milev MP, Leslie J, Saint-Dic D, Khan V, Hincapie AM, Day JO, McGavin L, Rowley C, Harlalka GV, Vancollie VE, Ahmed W, Lelliott CJ, Gul A, Yalcin B, Crosby AH, Sacher M, Baple EL. **Biallelic variants in *TRAPPC10* cause a microcephalic TRAPPopathy disorder in humans and mice**. *PLOS Genetics*. Mar 2022 18(3): e1010114.

Leslie JS*, **Rawlins LE***, Chioza BA, Olubodun OR, Salter CG, Fasham J, Jones HF, Cross HE, Lam S, Harlalka GV, Muggenthaler MMA, Crosby AH, Baple EL. ***MNS1* variant associated with situs inversus and male infertility.** *Eur J Hum Genet.* 2020 Jan;28(1):50-55.

Rawlins LE, Jones H, Wenger O, Aye M, Fasham J, Harlalka GV, Chioza BA, Miron A, Ellard S, Wakeling M, Crosby AH, Baple EL. **An Amish founder variant consolidates disruption of CEP55 as a cause of hydranencephaly and renal dysplasia**. *Eur J Hum Genet.* 2019 Apr;27(4):657-662.

Levitin MO, **Rawlins LE**, Sanchez-Andrade G, Arshad O, Collins SC, Sawiak SJ, Iffland PH, Andersson MHL, Bupp C, Cambridge EL, Coomber EL, Ellis I, Herkert JC, Ironfield H, Jory L, Kretz PF, Kant S, Neaverson A, Nibbeling E, Rowley C, Relton E, Sanderson M, Scott EM, Stewart H, Shuen AY, Schreiber J, Tuck L, Tonks J, Terkelsen T, van Ravenswaaij-Arts C, Vasudevan P, Wenger O, Wright M, Day A, Hunter A, Patel M, Lelliott CJ, Crino PB, Yalcin B, Crosby AH, Baple EL, Logan DW, Hurles ME, Gerety SS. **Mouse and cellular models of *KPTN*-related disorder implicate mTOR signalling in cognitive and overgrowth phenotypes.** *BioRxiv* Jul 2022, submitted to *Nat Commun.*

Ammous Z, **Rawlins LE**, Jones H, Leslie JS, Wenger O, Scott E, Deline J, Herr T, Evans R, Scheid A, Kennedy J, Chioza BA, Ames RM, Cross HE, Puffenberger EG, Harries L, Baple EL, Crosby AH. **A biallelic *SNIP1* Amish founder variant causes a recognizable neurodevelopmental disorder.** *PLOS Genetics.* Sep 2021 17(9): e1009803.

Khan S, **Rawlins LE**, Harlalka GV, Umair M, Ullah A, Shahzad S, Javed M, Baple EL, Crosby AH, Ahmad W, Gul A. **Homozygous variants in the *HEXB* and *MBOAT7* genes underlie neurological diseases in consanguineous families**. *BMC Med Genet.* 2019 20(1).

Fasham J, Lin S, Ghosh P, Radio FC, Farrow EG, Thiffault I, Kussman J, Zhou D, Hemming R, Zahka K, Chioza BA, **Rawlins LE**, Wenger OK, Gunning AC, Pizzi S, Onesimo R, Zampino G, Barker E, Osawa N, Rodriguez MC, Neuhann TM, Zackai EH, Keena B, Capasso J, Levin AV, Bhoj E, Li D, Hakonarson H, Wentzensen IM, Jackson A, Chandler KE, Coban-Akdemir ZH, Posey JE, Banka S, Lupski JR, Sheppard SE, Tartaglia M, Triggs-Raine B, Crosby AH, Baple EL. **Elucidating the clinical spectrum and molecular basis of HYAL2 deficiency.** *Genet Med.* 2021 Nov 30.

Yuan B, Neira J, Pehlivan D, Santiago-Sim T, Song X, Rosenfeld J, Posey JE, Patel V, Jin W, Adam MP, Baple EL, Dean J, Fong CT, Hickey SE, Hudgins L, Leon E, Madan-Khetarpal S, **Rawlins LE**, Rustad CF, Stray-Pedersen A, Tveten K, Wenger O, Diaz J, Jenkins L, Martin L, McGuire M, Pietryga M, Ramsdell L, Slattery L; DDD Study, Abid F, Bertuch AA, Grange D, Immken L, Schaaf CP, Van Esch H, Bi W, Cheung SW, Breman AM, Smith JL, Shaw C, Crosby AH, Eng C, Yang Y, Lupski JR, Xiao R, Liu P. **Clinical exome sequencing reveals locus heterogeneity and phenotypic variability of cohesinopathies.** Genet Med. 2019 Mar;21(3):663-675.

* joint first author

REFERENCES

1.   Sadler TW. Langman's Medical Embryology. 11th edition ed: Lippincott Williams and Wilkins; 2010.

2.   Sadler TW. Embryology of neural tube development. Am J Med Genet C Semin Med Genet. 2005;135C(1):2-8. doi: 10.1002/ajmg.c.30049. PubMed PMID: 15806586.

3.   Crossman AR, Neary D. Neuroanatomy: An illustrated colour text. 4th Edition ed: Churchill Livingstone; 2010.

4.   ten Donkelaar HJ, Lammens M, Wesseling P, Thijssen HO, Renier WO. Development and developmental disorders of the human cerebellum. J Neurol. 2003;250(9):1025-36. doi: 10.1007/s00415-003-0199-9. PubMed PMID: 14504962.

5.   Bear M, Connors B, Paradiso M. Neuroscience: Exploring the Brain. Fourth ed. Massachusetts, USA: Jones and Bartlett Learning; 2020.

6.   Rakic P. Evolution of the neocortex: a perspective from developmental biology. Nat Rev Neurosci. 2009;10(10):724-35. doi: 10.1038/nrn2719. PubMed PMID: 19763105; PubMed Central PMCID: 2913577.

7.   Rasika S, Passemard S, Verloes A, Gressens P, El Ghouzzi V. Golgipathies in Neurodevelopment: A New View of Old Defects. Dev Neurosci. 2018;40(5-6):396-416. Epub 2019/03/15. doi: 10.1159/000497035. PubMed PMID: 30878996.

8.   Noctor SC, Flint AC, Weissman TA, Wong WS, Clinton BK, Kriegstein AR. Dividing precursor cells of the embryonic cortical ventricular zone have morphological and molecular characteristics of radial glia. J Neurosci. 2002;22(8):3161-73. doi: 20026299. PubMed PMID: 11943818; PubMed Central PMCID: 6757532.

9.   Miyata T, Kawaguchi A, Saito K, Kawano M, Muto T, Ogawa M. Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells. Development. 2004;131(13):3133-45. Epub 20040602. doi: 10.1242/dev.01173. PubMed PMID: 15175243.

10.   Taverna E, Götz M, Huttner WB. The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. Annu Rev Cell Dev Biol. 2014;30:465-502. Epub 20140627. doi: 10.1146/annurev-cellbio-101011-155801. PubMed PMID: 25000993.

11.   Hansen DV, Lui JH, Parker PR, Kriegstein AR. Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature. 2010;464(7288):554-61. doi: 10.1038/nature08845. PubMed PMID: 20154730.

12.   Fernández V, Llinares-Benadero C, Borrell V. Cerebral cortex expansion and folding: what have we learned? EMBO J. 2016;35(10):1021-44. Epub

20160407. doi: 10.15252/embj.201593701. PubMed PMID: 27056680; PubMed Central PMCID: 4868950.

13.   deAzevedo LC, Fallet C, Moura-Neto V, Daumas-Duport C, Hedin-Pereira C, Lent R. Cortical radial glial cells in human fetuses: depth-correlated transformation into astrocytes. J Neurobiol. 2003;55(3):288-98. doi: 10.1002/neu.10205. PubMed PMID: 12717699.

14.   Martínez-Cerdeño V, Noctor SC, Kriegstein AR. The role of intermediate progenitor cells in the evolutionary expansion of the cerebral cortex. Cereb Cortex. 2006;16 Suppl 1:i152-61. doi: 10.1093/cercor/bhk017. PubMed PMID: 16766701.

15.   Nowakowski TJ, Pollen AA, Sandoval-Espinosa C, Kriegstein AR. Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral Cortex Development. Neuron. 2016;91(6):1219-27. doi: 10.1016/j.neuron.2016.09.005. PubMed PMID: 27657449; PubMed Central PMCID: 5087333.

16.   Tabata H, Nakajima K. Multipolar migration: the third mode of radial neuronal migration in the developing cerebral cortex. J Neurosci. 2003;23(31):9996-10001. PubMed PMID: 14602813; PubMed Central PMCID: 6740853.

17.   Nadarajah B, Brunstrom JE, Grutzendler J, Wong RO, Pearlman AL. Two modes of radial migration in early development of the cerebral cortex. Nat Neurosci. 2001;4(2):143-50. doi: 10.1038/83967. PubMed PMID: 11175874.

18.   Cooper JA. A mechanism for inside-out lamination in the neocortex. Trends Neurosci. 2008;31(3):113-9. Epub 20080205. doi: 10.1016/j.tins.2007.12.003. PubMed PMID: 18255163.

19.   Noctor SC, Flint AC, Weissman TA, Dammerman RS, Kriegstein AR. Neurons derived from radial glial cells establish radial units in neocortex. Nature. 2001;409(6821):714-20. doi: 10.1038/35055553. PubMed PMID: 11217860.

20.   Hirota Y, Nakajima K. Control of Neuronal Migration and Aggregation by Reelin Signaling in the Developing Cerebral Cortex. Front Cell Dev Biol. 2017;5:40. Epub 20170426. doi: 10.3389/fcell.2017.00040. PubMed PMID: 28507985; PubMed Central PMCID: 5410752.

21.   Elias LA, Kriegstein AR. Gap junctions: multifaceted regulators of embryonic cortical development. Trends Neurosci. 2008;31(5):243-50. Epub 20080409. doi: 10.1016/j.tins.2008.02.007. PubMed PMID: 18403031; PubMed Central PMCID: 2610634.

22.   Marín O, Valiente M, Ge X, Tsai LH. Guiding neuronal cell migrations. Cold Spring Harb Perspect Biol. 2010;2(2):a001834. doi: 10.1101/cshperspect.a001834. PubMed PMID: 20182622; PubMed Central PMCID: 2828271.

23.   Subramanian L, Calcagnotto ME, Paredes MF. Cortical Malformations: Lessons in Human Brain Development. Front Cell Neurosci. 2019;13:576. Epub 20200124. doi: 10.3389/fncel.2019.00576. PubMed PMID: 32038172; PubMed Central PMCID: 6993122.

24.   ten Donkelaar H, Lammens M, Aronica E, van Bokhoven H, Kamphuis-van Ulzen K, Hori A. Development and Developmental Disorders of the Cerebral Cortex. Clinical Neuroembryology. Berlin: Springer; 2014. p. 523-642.

25.   Rakic P, Lombroso PJ. Development of the cerebral cortex: I. Forming the cortical structure. J Am Acad Child Adolesc Psychiatry. 1998;37(1):116-7. doi: 10.1097/00004583-199801000-00026. PubMed PMID: 9444908.

26.   Raybaud C, Ahmad T, Rastegar N, Shroff M, Al Nassar M. The premature brain: developmental and lesional anatomy. Neuroradiology. 2013;55 Suppl 2:23-40. Epub 20130707. doi: 10.1007/s00234-013-1231-0. PubMed PMID: 23832006.

27.   Herschkowitz N, Kagan J, Zilles K. Neurobiological bases of behavioral development in the second year. Neuropediatrics. 1999;30(5):221-30. doi: 10.1055/s-2007-973495. PubMed PMID: 10598832.

28.   Barres BA. The mystery and magic of glia: a perspective on their roles in health and disease. Neuron. 2008;60(3):430-40. doi: 10.1016/j.neuron.2008.10.013. PubMed PMID: 18995817.

29.   Takahashi T, Nowakowski RS, Caviness VS. The cell cycle of the pseudostratified ventricular epithelium of the embryonic murine cerebral wall. J Neurosci. 1995;15(9):6046-57. PubMed PMID: 7666188; PubMed Central PMCID: 6577667.

30.   Tidyman WE, Rauen KA. The RASopathies: developmental syndromes of Ras/MAPK pathway dysregulation. Curr Opin Genet Dev. 2009;19(3):230-6. Epub 20090519. doi: 10.1016/j.gde.2009.04.001. PubMed PMID: 19467855; PubMed Central PMCID: 2743116.

31.   Tee AR, Sampson JR, Pal DK, Bateman JM. The role of mTOR signalling in neurogenesis, insights from tuberous sclerosis complex. Semin Cell Dev Biol. 2016;52:12-20. Epub 20160202. doi: 10.1016/j.semcdb.2016.01.040. PubMed PMID: 26849906; PubMed Central PMCID: 6379054.

32.   Marin Navarro A, Pronk RJ, van der Geest AT, Oliynyk G, Nordgren A, Arsenian-Henriksson M, et al. p53 controls genomic stability and temporal differentiation of human neural stem cells and affects neural organization in human brain organoids. Cell Death Dis. 2020;11(1):52. Epub 20200123. doi: 10.1038/s41419-019-2208-7. PubMed PMID: 31974372; PubMed Central PMCID: 6978389.

33.   Haydar TF, Ang E, Rakic P. Mitotic spindle rotation and mode of cell division in the developing telencephalon. Proc Natl Acad Sci U S A. 2003;100(5):2890-5. Epub 20030214. doi: 10.1073/pnas.0437969100. PubMed PMID: 12589023; PubMed Central PMCID: 151436.

34.   Łukasik P, Załuski M, Gutowska I. Cyclin-Dependent Kinases (CDK) and Their Role in Diseases Development-Review. Int J Mol Sci. 2021;22(6). Epub 20210313. doi: 10.3390/ijms22062935. PubMed PMID: 33805800; PubMed Central PMCID: 7998717.

35.   Nicolas CS, Amici M, Bortolotto ZA, Doherty A, Csaba Z, Fafouri A, et al. The role of JAK-STAT signaling within the CNS. JAKSTAT. 2013;2(1):e22925. doi: 10.4161/jkst.22925. PubMed PMID: 24058789; PubMed Central PMCID: 3670265.

36.   Kuan CY, Roth KA, Flavell RA, Rakic P. Mechanisms of programmed cell death in the developing brain. Trends Neurosci. 2000;23(7):291-7. doi: 10.1016/s0166-2236(00)01581-2. PubMed PMID: 10856938.

37.   Takei N, Nawa H. mTOR signaling and its roles in normal and abnormal brain development. Front Mol Neurosci. 2014;7:28. Epub 20140423. doi: 10.3389/fnmol.2014.00028. PubMed PMID: 24795562; PubMed Central PMCID: 4005960.

38.   Dias JM, Alekseenko Z, Applequist JM, Ericson J. Tgfβ signaling regulates temporal neurogenesis and potency of neural stem cells in the CNS. Neuron. 2014;84(5):927-39. Epub 20141113. doi: 10.1016/j.neuron.2014.10.033. PubMed PMID: 25467979.

39.   Habas R, Dawid IB. Dishevelled and Wnt signaling: is the nucleus the final frontier? J Biol. 2005;4(1):2. Epub 20050217. doi: 10.1186/jbiol22. PubMed PMID: 15720723; PubMed Central PMCID: 551522.

40.   Frade JM. Interkinetic nuclear movement in the vertebrate neuroepithelium: encounters with an old acquaintance. Prog Brain Res. 2002;136:67-71. doi: 10.1016/s0079-6123(02)36007-2. PubMed PMID: 12143404.

41.   Del Bene F, Wehman AM, Link BA, Baier H. Regulation of neurogenesis by interkinetic nuclear migration through an apical-basal notch gradient. Cell. 2008;134(6):1055-65. doi: 10.1016/j.cell.2008.07.017. PubMed PMID: 18805097; PubMed Central PMCID: 2628487.

42.   Willaredt MA, Tasouri E, Tucker KL. Primary cilia and forebrain development. Mech Dev. 2013;130(6-8):373-80. Epub 20121022. doi: 10.1016/j.mod.2012.10.003. PubMed PMID: 23085524.

43.   Wheway G, Nazlamova L, Hancock JT. Signaling through the Primary Cilium. Front Cell Dev Biol. 2018;6:8. Epub 20180208. doi: 10.3389/fcell.2018.00008. PubMed PMID: 29473038; PubMed Central PMCID: 5809511.

44.   Sheen VL, Ganesh VS, Topcu M, Sebire G, Bodell A, Hill RS, et al. Mutations in ARFGEF2 implicate vesicle trafficking in neural progenitor proliferation and migration in the human cerebral cortex. Nat Genet. 2004;36(1):69-76. Epub 20031130. doi: 10.1038/ng1276. PubMed PMID: 14647276.

45.   Raghu P, Joseph A, Krishnan H, Singh P, Saha S. Phosphoinositides: Regulators of Nervous System Function in Health and Disease. Front Mol Neurosci. 2019;12:208. Epub 20190823. doi: 10.3389/fnmol.2019.00208. PubMed PMID: 31507376; PubMed Central PMCID: 6716428.

46.   Lehtinen MK, Zappaterra MW, Chen X, Yang YJ, Hill AD, Lun M, et al. The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. Neuron. 2011;69(5):893-905. doi: 10.1016/j.neuron.2011.01.023. PubMed PMID: 21382550; PubMed Central PMCID: 3085909.

47.   Ma X, Qiu S. Control of cortical synapse development and plasticity by MET receptor tyrosine kinase, a genetic risk factor for autism. J Neurosci Res. 2020;98(11):2115-29. Epub 20191119. doi: 10.1002/jnr.24542. PubMed PMID: 31746037; PubMed Central PMCID: 7234907.

48.   Graus-Porta D, Blaess S, Senften M, Littlewood-Evans A, Damsky C, Huang Z, et al. Beta1-class integrins regulate the development of laminae and folia in the cerebral and cerebellar cortex. Neuron. 2001;31(3):367-79. doi: 10.1016/s0896-6273(01)00374-9. PubMed PMID: 11516395.

49.   Barresi R, Campbell KP. Dystroglycan: from biosynthesis to pathogenesis of human disease. J Cell Sci. 2006;119(Pt 2):199-207. doi: 10.1242/jcs.02814. PubMed PMID: 16410545.

50.   Gripp KW, Slavotinek AM, Hall JG, Allanson JE. Handbook of physical measurements. 3rd edition ed: Oxford; 2013. p. 77-85.

51.   Woods CG, Parker A. Investigating microcephaly. Arch Dis Child. 2013;98(9):707-13. Epub 20130628. doi: 10.1136/archdischild-2012-302882. PubMed PMID: 23814088.

52.   Group WMGRS. Assessment of differences in linear growth among populations in the WHO Multicentre Growth Reference Study. Acta Paediatr Suppl. 2006;450:56-65. doi: 10.1111/j.1651-2227.2006.tb02376.x. PubMed PMID: 16817679.

53.   Cole TJ. The development of growth references and growth charts. Ann Hum Biol. 2012;39(5):382-94. Epub 20120711. doi: 10.3109/03014460.2012.694475. PubMed PMID: 22780429; PubMed Central PMCID: 3920659.

54.   Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med. 1998;17(4):407-29. PubMed PMID: 9496720.

55.   Group WMGRS. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl. 2006;450:76-85. doi: 10.1111/j.1651-2227.2006.tb02378.x. PubMed PMID: 16817681.

56.   Natale V, Rajagopalan A. Worldwide variation in human growth and the World Health Organization growth standards: a systematic review. BMJ Open.

2014;4(1):e003735. Epub 20140108. doi: 10.1136/bmjopen-2013-003735. PubMed PMID: 24401723; PubMed Central PMCID: 3902406.

57. Bale SJ, Amos CI, Parry DM, Bale AE. Relationship between head circumference and height in normal adults and in the nevoid basal cell carcinoma syndrome and neurofibromatosis type I. Am J Med Genet. 1991;40(2):206-10. doi: 10.1002/ajmg.1320400217. PubMed PMID: 1910262.

58. Bushby KM, Cole T, Matthews JN, Goodship JA. Centiles for adult head circumference. Arch Dis Child. 1992;67(10):1286-7. doi: 10.1136/adc.67.10.1286. PubMed PMID: 1444530; PubMed Central PMCID: 1793909.

59. Weaver DD, Christian JC. Familial variation of head size and adjustment for parental head circumference. J Pediatr. 1980;96(6):990-4. doi: 10.1016/s0022-3476(80)80623-8. PubMed PMID: 7189556.

60. Mitchison HM, Valente EM. Motile and non-motile cilia in human pathology: from function to phenotypes. J Pathol. 2017;241(2):294-309. doi: 10.1002/path.4843. PubMed PMID: 27859258.

61. Ceccaldi R, Sarangi P, D'Andrea AD. The Fanconi anaemia pathway: new players and new functions. Nat Rev Mol Cell Biol. 2016;17(6):337-49. Epub 20160505. doi: 10.1038/nrm.2016.48. PubMed PMID: 27145721.

62. Robertson SP. Filamin A: phenotypic diversity. Curr Opin Genet Dev. 2005;15(3):301-7. doi: 10.1016/j.gde.2005.04.001. PubMed PMID: 15917206.

63. Lappalainen T, Scott AJ, Brandt M, Hall IM. Genomic Analysis in the Age of Human Genome Sequencing. Cell. 2019;177(1):70-84. doi: 10.1016/j.cell.2019.02.032. PubMed PMID: 30901550; PubMed Central PMCID: 6532068.

64. Posey JE. Genome sequencing and implications for rare disorders. Orphanet J Rare Dis. 2019;14(1):153. Epub 20190624. doi: 10.1186/s13023-019-1127-0. PubMed PMID: 31234920; PubMed Central PMCID: 6591893.

65. Vissers LE, Gilissen C, Veltman JA. Genetic studies in intellectual disability and related disorders. Nat Rev Genet. 2016;17(1):9-18. Epub 20151027. doi: 10.1038/nrg3999. PubMed PMID: 26503795.

66. Musante L, Ropers HH. Genetics of recessive cognitive disorders. Trends Genet. 2014;30(1):32-9. Epub 20131028. doi: 10.1016/j.tig.2013.09.008. PubMed PMID: 24176302.

67. Abuelo D. Microcephaly syndromes. Semin Pediatr Neurol. 2007;14(3):118-27. doi: 10.1016/j.spen.2007.07.003. PubMed PMID: 17980308.

68. McKusick-Nathans Institute of Genetic Medicine JHU. Online Mendelian Inheritance in Man Baltimore, MD2019 [cited 25 January 2019]. Available from: https://www.omim.org/.

69.   Passemard S, Kaindl AM, Verloes A. Microcephaly. Handb Clin Neurol. 2013;111:129-41. doi: 10.1016/B978-0-444-52891-9.00013-0. PubMed PMID: 23622158.

70.   Battaglia A, Carey JC, South ST. Wolf-Hirschhorn syndrome: A review and update. Am J Med Genet C Semin Med Genet. 2015;169(3):216-23. Epub 20150804. doi: 10.1002/ajmg.c.31449. PubMed PMID: 26239400.

71.   Pober BR. Williams-Beuren syndrome. N Engl J Med. 2010;362(3):239-52. doi: 10.1056/NEJMra0903074. PubMed PMID: 20089974.

72.   Bond J, Roberts E, Mochida GH, Hampshire DJ, Scott S, Askham JM, et al. ASPM is a major determinant of cerebral cortical size. Nat Genet. 2002;32(2):316-20. Epub 2002/09/23. doi: 10.1038/ng995. PubMed PMID: 12355089.

73.   Bilgüvar K, Oztürk AK, Louvi A, Kwan KY, Choi M, Tatli B, et al. Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. Nature. 2010;467(7312):207-10. Epub 2010/08/22. doi: 10.1038/nature09327. PubMed PMID: 20729831; PubMed Central PMCID: 3129007.

74.   Bond J, Roberts E, Springell K, Lizarraga SB, Lizarraga S, Scott S, et al. A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. Nat Genet. 2005;37(4):353-5. Epub 2005/03/27. doi: 10.1038/ng1539. PubMed PMID: 15793586.

75.   Li H, Bielas SL, Zaki MS, Ismail S, Farfara D, Um K, et al. Biallelic Mutations in Citron Kinase Link Mitotic Cytokinesis to Human Primary Microcephaly. Am J Hum Genet. 2016;99(2):501-10. Epub 20160721. doi: 10.1016/j.ajhg.2016.07.004. PubMed PMID: 27453578; PubMed Central PMCID: 4974110.

76.   Moawia A, Shaheen R, Rasool S, Waseem SS, Ewida N, Budde B, et al. Mutations of KIF14 cause primary microcephaly by impairing cytokinesis. Ann Neurol. 2017;82(4):562-77. Epub 20171014. doi: 10.1002/ana.25044. PubMed PMID: 28892560.

77.   Kadir R, Harel T, Markus B, Perez Y, Bakhrat A, Cohen I, et al. ALFY-Controlled DVL3 Autophagy Regulates Wnt Signaling, Determining Human Brain Size. PLoS Genet. 2016;12(3):e1005919. Epub 20160323. doi: 10.1371/journal.pgen.1005919. PubMed PMID: 27008544; PubMed Central PMCID: 4805177.

78.   Yang YJ, Baltus AE, Mathew RS, Murphy EA, Evrony GD, Gonzalez DM, et al. Microcephaly gene links trithorax and REST/NRSF to control neural stem cell proliferation and differentiation. Cell. 2012;151(5):1097-112. doi: 10.1016/j.cell.2012.10.043. PubMed PMID: 23178126; PubMed Central PMCID: 3567437.

79.   He H, Liyanarachchi S, Akagi K, Nagy R, Li J, Dietrich RC, et al. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental

disorder MOPD I. Science. 2011;332(6026):238-40. doi: 10.1126/science.1200587. PubMed PMID: 21474760; PubMed Central PMCID: 3380448.

80. Griffith E, Walker S, Martin CA, Vagnarelli P, Stiff T, Vernay B, et al. Mutations in pericentrin cause Seckel syndrome with defective ATR-dependent DNA damage signaling. Nat Genet. 2008;40(2):232-6. Epub 20071223. doi: 10.1038/ng.2007.80. PubMed PMID: 18157127; PubMed Central PMCID: 2397541.

81. Klingseisen A, Jackson AP. Mechanisms and pathways of growth failure in primordial dwarfism. Genes Dev. 2011;25(19):2011-24. doi: 10.1101/gad.169037. PubMed PMID: 21979914; PubMed Central PMCID: 3197200.

82. Bicknell LS, Bongers EM, Leitch A, Brown S, Schoots J, Harley ME, et al. Mutations in the pre-replication complex cause Meier-Gorlin syndrome. Nat Genet. 2011;43(4):356-9. Epub 20110227. doi: 10.1038/ng.775. PubMed PMID: 21358632; PubMed Central PMCID: 3068194.

83. van der Burgt I, Chrzanowska KH, Smeets D, Weemaes C. Nijmegen breakage syndrome. J Med Genet. 1996;33(2):153-6. doi: 10.1136/jmg.33.2.153. PubMed PMID: 8929954; PubMed Central PMCID: 1051843.

84. German J. Bloom's syndrome. I. Genetical and clinical observations in the first twenty-seven patients. Am J Hum Genet. 1969;21(2):196-227. PubMed PMID: 5770175; PubMed Central PMCID: 1706430.

85. Shen J, Gilmore EC, Marshall CA, Haddadin M, Reynolds JJ, Eyaid W, et al. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat Genet. 2010;42(3):245-9. Epub 20100131. doi: 10.1038/ng.526. PubMed PMID: 20118933; PubMed Central PMCID: 2835984.

86. Waters AM, Asfahani R, Carroll P, Bicknell L, Lescai F, Bright A, et al. The kinetochore protein, CENPF, is mutated in human ciliopathy and microcephaly phenotypes. J Med Genet. 2015;52(3):147-56. Epub 20150106. doi: 10.1136/jmedgenet-2014-102691. PubMed PMID: 25564561; PubMed Central PMCID: 4345935.

87. Hildebrandt F, Benzing T, Katsanis N. Ciliopathies. N Engl J Med. 2011;364(16):1533-43. doi: 10.1056/NEJMra1010172. PubMed PMID: 21506742; PubMed Central PMCID: 3640822.

88. Hu H, Xiao X, Li S, Jia X, Guo X, Zhang Q. KIF11 mutations are a common cause of autosomal dominant familial exudative vitreoretinopathy. Br J Ophthalmol. 2016;100(2):278-83. Epub 20151015. doi: 10.1136/bjophthalmol-2015-306878. PubMed PMID: 26472404.

89. Martin CA, Ahmad I, Klingseisen A, Hussain MS, Bicknell LS, Leitch A, et al. Mutations in PLK4, encoding a master regulator of centriole biogenesis, cause microcephaly, growth failure and retinopathy. Nat Genet. 2014;46(12):1283-92.

Epub 20141026. doi: 10.1038/ng.3122. PubMed PMID: 25344692; PubMed Central PMCID: 4676084.

90.   Tabatabaie L, Klomp LW, Berger R, de Koning TJ. L-serine synthesis in the central nervous system: a review on serine deficiency disorders. Mol Genet Metab. 2010;99(3):256-62. Epub 20091025. doi: 10.1016/j.ymgme.2009.10.012. PubMed PMID: 19963421.

91.   Rosenberg MJ, Agarwala R, Bouffard G, Davis J, Fiermonte G, Hilliard MS, et al. Mutant deoxynucleotide carrier is associated with congenital microcephaly. Nat Genet. 2002;32(1):175-9. Epub 20020819. doi: 10.1038/ng948. PubMed PMID: 12185364.

92.   Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. Nat Genet. 2004;36(6):636-41. Epub 20040516. doi: 10.1038/ng1363. PubMed PMID: 15146185.

93.   Dobyns WB, Stratton RF, Parke JT, Greenberg F, Nussbaum RL, Ledbetter DH. Miller-Dieker syndrome: lissencephaly and monosomy 17p. J Pediatr. 1983;102(4):552-8. doi: 10.1016/s0022-3476(83)80183-8. PubMed PMID: 6834189.

94.   Jaeken J. Congenital disorders of glycosylation. Handb Clin Neurol. 2013;113:1737-43. doi: 10.1016/B978-0-444-59565-2.00044-7. PubMed PMID: 23622397.

95.   Firth H, Hurst J. Oxford Desk Reference: Clinical Genetics and Genomics. 2 ed. Oxford, UK: Oxford University Press; 2017.

96.   Aubourg P, Wanders R. Peroxisomal disorders. Handb Clin Neurol. 2013;113:1593-609. doi: 10.1016/B978-0-444-59565-2.00028-9. PubMed PMID: 23622381.

97.   Hedlund GL, Longo N, Pasquali M. Glutaric acidemia type 1. Am J Med Genet C Semin Med Genet. 2006;142C(2):86-94. doi: 10.1002/ajmg.c.30088. PubMed PMID: 16602100; PubMed Central PMCID: 2556991.

98.   Ozand PT, Gascon GG. Organic acidurias: a review. Part 1. J Child Neurol. 1991;6(3):196-219. doi: 10.1177/088307389100600302. PubMed PMID: 1875022.

99.   Pascual JM, Wang D, Lecumberri B, Yang H, Mao X, Yang R, et al. GLUT1 deficiency and other glucose transporter diseases. Eur J Endocrinol. 2004;150(5):627-33. doi: 10.1530/eje.0.1500627. PubMed PMID: 15132717.

100.   MENKES JH, ALTER M, STEIGLEDER GK, WEAKLEY DR, SUNG JH. A sex-linked recessive disorder with retardation of growth, peculiar hair, and focal cerebral and cerebellar degeneration. Pediatrics. 1962;29:764-79. PubMed PMID: 14472668.

101.   Stephenson JB. Aicardi-Goutières syndrome (AGS). Eur J Paediatr Neurol. 2008;12(5):355-8. Epub 20080314. doi: 10.1016/j.ejpn.2007.11.010. PubMed PMID: 18343173.

102.   Nance MA, Berry SA. Cockayne syndrome: review of 140 cases. Am J Med Genet. 1992;42(1):68-84. doi: 10.1002/ajmg.1320420115. PubMed PMID: 1308368.

103.   Hagberg B, Aicardi J, Dias K, Ramos O. A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol. 1983;14(4):471-9. doi: 10.1002/ana.410140412. PubMed PMID: 6638958.

104.   Kolehmainen J, Black GC, Saarinen A, Chandler K, Clayton-Smith J, Träskelin AL, et al. Cohen syndrome is caused by mutations in a novel gene, COH1, encoding a transmembrane protein with a presumed role in vesicle-mediated sorting and intracellular protein transport. Am J Hum Genet. 2003;72(6):1359-69. Epub 20030502. doi: 10.1086/375454. PubMed PMID: 12730828; PubMed Central PMCID: 1180298.

105.   Sacher M, Shahrzad N, Kamel H, Milev MP. TRAPPopathies: An emerging set of disorders linked to variations in the genes encoding transport protein particle (TRAPP)-associated proteins. Traffic. 2019;20(1):5-26. Epub 2018/09/24. doi: 10.1111/tra.12615. PubMed PMID: 30152084.

106.   Passemard S, Perez F, Colin-Lemesre E, Rasika S, Gressens P, El Ghouzzi V. Golgi trafficking defects in postnatal microcephaly: The evidence for "Golgipathies". Prog Neurobiol. 2017;153:46-63. Epub 2017/04/02. doi: 10.1016/j.pneurobio.2017.03.007. PubMed PMID: 28377289.

107.   Romaniello R, Arrigoni F, Fry AE, Bassi MT, Rees MI, Borgatti R, et al. Tubulin genes and malformations of cortical development. Eur J Med Genet. 2018;61(12):744-54. Epub 20180717. doi: 10.1016/j.ejmg.2018.07.012. PubMed PMID: 30016746.

108.   Gleeson JG, Allen KM, Fox JW, Lamperti ED, Berkovic S, Scheffer I, et al. Doublecortin, a brain-specific gene mutated in human X-linked lissencephaly and double cortex syndrome, encodes a putative signaling protein. Cell. 1998;92(1):63-72. doi: 10.1016/s0092-8674(00)80899-5. PubMed PMID: 9489700.

109.   Poirier K, Keays DA, Francis F, Saillour Y, Bahi N, Manouvrier S, et al. Large spectrum of lissencephaly and pachygyria phenotypes resulting from de novo missense mutations in tubulin alpha 1A (TUBA1A). Hum Mutat. 2007;28(11):1055-64. doi: 10.1002/humu.20572. PubMed PMID: 17584854.

110.   Guven A, Gunduz A, Bozoglu TM, Yalcinkaya C, Tolun A. Novel NDE1 homozygous mutation resulting in microhydranencephaly and not microlyssencephaly. Neurogenetics. 2012;13(3):189-94. Epub 20120415. doi: 10.1007/s10048-012-0326-9. PubMed PMID: 22526350.

111.   Jaglin XH, Poirier K, Saillour Y, Buhler E, Tian G, Bahi-Buisson N, et al. Mutations in the beta-tubulin gene TUBB2B result in asymmetrical polymicrogyria. Nat Genet. 2009;41(6):746-52. Epub 20090524. doi: 10.1038/ng.380. PubMed PMID: 19465910; PubMed Central PMCID: 2883584.

112.   Zhang X, Ling J, Barcia G, Jing L, Wu J, Barry BJ, et al. Mutations in QARS, encoding glutaminyl-tRNA synthetase, cause progressive microcephaly, cerebral-cerebellar atrophy, and intractable seizures. Am J Hum Genet. 2014;94(4):547-58. Epub 20140320. doi: 10.1016/j.ajhg.2014.03.003. PubMed PMID: 24656866; PubMed Central PMCID: 3980424.

113.   Nakayama T, Wu J, Galvin-Parton P, Weiss J, Andriola MR, Hill RS, et al. Deficient activity of alanyl-tRNA synthetase underlies an autosomal recessive syndrome of progressive microcephaly, hypomyelination, and epileptic encephalopathy. Hum Mutat. 2017;38(10):1348-54. Epub 20170623. doi: 10.1002/humu.23250. PubMed PMID: 28493438; PubMed Central PMCID: 5599341.

114.   Kuo ME, Theil AF, Kievit A, Malicdan MC, Introne WJ, Christian T, et al. Cysteinyl-tRNA Synthetase Mutations Cause a Multi-System, Recessive Disease That Includes Microcephaly, Developmental Delay, and Brittle Hair and Nails. Am J Hum Genet. 2019;104(3):520-9. Epub 20190226. doi: 10.1016/j.ajhg.2019.01.006. PubMed PMID: 30824121; PubMed Central PMCID: 6407526.

115.   Bond J, Woods CG. Cytoskeletal genes regulating brain size. Curr Opin Cell Biol. 2006;18(1):95-101. Epub 20051206. doi: 10.1016/j.ceb.2005.11.004. PubMed PMID: 16337370.

116.   Jayaraman D, Bae BI, Walsh CA. The Genetics of Primary Microcephaly. Annu Rev Genomics Hum Genet. 2018;19:177-200. Epub 20180523. doi: 10.1146/annurev-genom-083117-021441. PubMed PMID: 29799801.

117.   Gilmore EC, Walsh CA. Genetic causes of microcephaly and lessons for neuronal development. Wiley Interdiscip Rev Dev Biol. 2013;2(4):461-78. Epub 20121004. doi: 10.1002/wdev.89. PubMed PMID: 24014418; PubMed Central PMCID: 3767923.

118.   Murray JE, Jackson AP. Exploring microcephaly and human brain evolution. Dev Med Child Neurol. 2012;54(7):580-1. Epub 2012/05/09. doi: 10.1111/j.1469-8749.2012.04330.x. PubMed PMID: 22571810.

119.   Zaqout S, Kaindl AM. Autosomal Recessive Primary Microcephaly: Not Just a Small Brain. Front Cell Dev Biol. 2021;9:784700. Epub 20220117. doi: 10.3389/fcell.2021.784700. PubMed PMID: 35111754; PubMed Central PMCID: 8802810.

120.   Passemard S, Titomanlio L, Elmaleh M, Afenjar A, Alessandri JL, Andria G, et al. Expanding the clinical and neuroradiologic phenotype of primary microcephaly due to ASPM mutations. Neurology. 2009;73(12):962-9. doi: 10.1212/WNL.0b013e3181b8799a. PubMed PMID: 19770472.

121. Manzini MC, Walsh CA. What disorders of cortical development tell us about the cortex: one plus one does not always make two. Curr Opin Genet Dev. 2011;21(3):333-9. Epub 20110201. doi: 10.1016/j.gde.2011.01.006. PubMed PMID: 21288712; PubMed Central PMCID: 3139684.

122. Buchman JJ, Durak O, Tsai LH. ASPM regulates Wnt signaling pathway activity in the developing brain. Genes Dev. 2011;25(18):1909-14. doi: 10.1101/gad.16830211. PubMed PMID: 21937711; PubMed Central PMCID: 3185963.

123. Martin AR, Williams E, Foulger RE, Leigh S, Daugherty LC, Niblock O, et al. PanelApp crowdsources expert knowledge to establish consensus diagnostic gene panels. Nat Genet. 2019;51(11):1560-5. doi: 10.1038/s41588-019-0528-2. PubMed PMID: 31676867.

124. DeMyer W. Megalencephaly: types, clinical syndromes, and management. Pediatr Neurol. 1986;2(6):321-8. doi: 10.1016/0887-8994(86)90072-x. PubMed PMID: 3334205.

125. Pirozzi F, Nelson B, Mirzaa G. From microcephaly to megalencephaly: determinants of brain size. Dialogues Clin Neurosci. 2018;20(4):267-82. PubMed PMID: 30936767; PubMed Central PMCID: 6436952.

126. Winden KD, Yuskaitis CJ, Poduri A. Megalencephaly and Macrocephaly. Semin Neurol. 2015;35(3):277-87. Epub 20150610. doi: 10.1055/s-0035-1552622. PubMed PMID: 26060907.

127. Gürbüz BB, Yılmaz DY, Coşkun T, Tokatlı A, Dursun A, Sivri HS. Glutaric aciduria type 1: Genetic and phenotypic spectrum in 53 patients. Eur J Med Genet. 2020;63(11):104032. Epub 20200807. doi: 10.1016/j.ejmg.2020.104032. PubMed PMID: 32777384.

128. Steenweg ME, Jakobs C, Errami A, van Dooren SJ, Adeva Bartolomé MT, Aerssens P, et al. An overview of L-2-hydroxyglutarate dehydrogenase gene (L2HGDH) variants: a genotype-phenotype study. Hum Mutat. 2010;31(4):380-90. doi: 10.1002/humu.21197. PubMed PMID: 20052767.

129. Kranendijk M, Struys EA, Gibson KM, Wickenhagen WV, Abdenur JE, Buechner J, et al. Evidence for genetic heterogeneity in D-2-hydroxyglutaric aciduria. Hum Mutat. 2010;31(3):279-83. doi: 10.1002/humu.21186. PubMed PMID: 20020533.

130. Renaud DL. Leukoencephalopathies associated with macrocephaly. Semin Neurol. 2012;32(1):34-41. Epub 20120315. doi: 10.1055/s-0032-1306384. PubMed PMID: 22422204.

131. Matalon R, Kaul R, Casanova J, Michals K, Johnson A, Rapin I, et al. SSIEM Award. Aspartoacylase deficiency: the enzyme defect in Canavan disease. J Inherit Metab Dis. 1989;12 Suppl 2:329-31. doi: 10.1007/BF03335413. PubMed PMID: 2512436.

132. Gorospe JR, Naidu S, Johnson AB, Puri V, Raymond GV, Jenkins SD, et al. Molecular findings in symptomatic and pre-symptomatic Alexander disease patients. Neurology. 2002;58(10):1494-500. doi: 10.1212/wnl.58.10.1494. PubMed PMID: 12034785.

133. Bugiani M, Boor I, Powers JM, Scheper GC, van der Knaap MS. Leukoencephalopathy with vanishing white matter: a review. J Neuropathol Exp Neurol. 2010;69(10):987-96. doi: 10.1097/NEN.0b013e3181f2eafa. PubMed PMID: 20838246.

134. López-Hernández T, Ridder MC, Montolio M, Capdevila-Nortes X, Polder E, Sirisi S, et al. Mutant GlialCAM causes megalencephalic leukoencephalopathy with subcortical cysts, benign familial macrocephaly, and macrocephaly with retardation and autism. Am J Hum Genet. 2011;88(4):422-32. Epub 20110317. doi: 10.1016/j.ajhg.2011.02.009. PubMed PMID: 21419380; PubMed Central PMCID: 3071909.

135. Suzuki K, Rapin I, Suzuki Y, Ishii N. Juvenile GM2-gangliosidosis. Clinical variant of Tay-Sachs disease or a new disease. Neurology. 1970;20(2):190-204. doi: 10.1212/wnl.20.2.190. PubMed PMID: 5460705.

136. Sandhoff K, Andreae U, Jatzkewitz H. Deficient hexosaminidase activity in an exceptional case of Tay-Sachs disease with additional storage of kidney globoside in visceral organs. Pathol Eur. 1968;3(2):278-85. PubMed PMID: 5688464.

137. D'Avanzo F, Rigon L, Zanetti A, Tomanin R. Mucopolysaccharidosis Type II: One Hundred Years of Research, Diagnosis, and Treatment. Int J Mol Sci. 2020;21(4). Epub 20200213. doi: 10.3390/ijms21041258. PubMed PMID: 32070051; PubMed Central PMCID: 7072947.

138. Cleary MA, Wraith JE. The presenting features of mucopolysaccharidosis type IH (Hurler syndrome). Acta Paediatr. 1995;84(3):337-9. doi: 10.1111/j.1651-2227.1995.tb13640.x. PubMed PMID: 7780260.

139. Tomanin R, Karageorgos L, Zanetti A, Al-Sayed M, Bailey M, Miller N, et al. Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. Hum Mutat. 2018;39(12):1788-802. Epub 20180917. doi: 10.1002/humu.23613. PubMed PMID: 30118150; PubMed Central PMCID: 6282714.

140. Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al. GeneReviews. 1993.

141. Tappino B, Biancheri R, Mort M, Regis S, Corsolini F, Rossi A, et al. Identification and characterization of 15 novel GALC gene mutations causing Krabbe disease. Hum Mutat. 2010;31(12):E1894-914. doi: 10.1002/humu.21367. PubMed PMID: 20886637; PubMed Central PMCID: 3052420.

142. Rivière JB, Mirzaa GM, O'Roak BJ, Beddaoui M, Alcantara D, Conway RL, et al. De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. Nat Genet.

2012;44(8):934-40. Epub 20120624. doi: 10.1038/ng.2331. PubMed PMID: 22729224; PubMed Central PMCID: 3408813.

143. Orlova KA, Crino PB. The tuberous sclerosis complex. Ann N Y Acad Sci. 2010;1184:87-105. doi: 10.1111/j.1749-6632.2009.05117.x. PubMed PMID: 20146692; PubMed Central PMCID: 2892799.

144. Crino PB. Focal brain malformations: seizures, signaling, sequencing. Epilepsia. 2009;50 Suppl 9:3-8. doi: 10.1111/j.1528-1167.2009.02289.x. PubMed PMID: 19761448.

145. Tatton-Brown K, Douglas J, Coleman K, Baujat G, Cole TR, Das S, et al. Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. Am J Hum Genet. 2005;77(2):193-204. Epub 20050607. doi: 10.1086/432082. PubMed PMID: 15942875; PubMed Central PMCID: 1224542.

146. Crino PB. mTOR: A pathogenic signaling pathway in developmental brain malformations. Trends Mol Med. 2011;17(12):734-42. Epub 2011/09/02. doi: 10.1016/j.molmed.2011.07.008. PubMed PMID: 21890410.

147. Hobert JA, Eng C. PTEN hamartoma tumor syndrome: an overview. Genet Med. 2009;11(10):687-94. doi: 10.1097/GIM.0b013e3181ac9aea. PubMed PMID: 19668082.

148. Smith LD, Saunders CJ, Dinwiddie DL, Atherton AM, Miller NA, Soden SE, et al. Exome sequencing reveals de novo germline mutation of mammalian target of rapamycin (MTOR) in a patient with megalencephaly and intractable seizures. J Genomes Exomes. 2013;2:10.

149. Roberts AE, Allanson JE, Tartaglia M, Gelb BD. Noonan syndrome. Lancet. 2013;381(9863):333-42. Epub 20130110. doi: 10.1016/S0140-6736(12)61023-X. PubMed PMID: 23312968; PubMed Central PMCID: 4267483.

150. Ferner RE, Huson SM, Thomas N, Moss C, Willshaw H, Evans DG, et al. Guidelines for the diagnosis and management of individuals with neurofibromatosis 1. J Med Genet. 2007;44(2):81-8. Epub 20061114. doi: 10.1136/jmg.2006.045906. PubMed PMID: 17105749; PubMed Central PMCID: 2598063.

151. Roberts A, Allanson J, Jadico SK, Kavamura MI, Noonan J, Opitz JM, et al. The cardiofaciocutaneous syndrome. J Med Genet. 2006;43(11):833-42. Epub 20060706. doi: 10.1136/jmg.2006.042796. PubMed PMID: 16825433; PubMed Central PMCID: 2563180.

152. Gripp KW, Morse LA, Axelrad M, Chatfield KC, Chidekel A, Dobyns W, et al. Costello syndrome: Clinical phenotype, genotype, and management guidelines. Am J Med Genet A. 2019;179(9):1725-44. Epub 20190620. doi: 10.1002/ajmg.a.61270. PubMed PMID: 31222966; PubMed Central PMCID: 8238015.

153. Klein SD, Nguyen DC, Bhakta V, Wong D, Chang VY, Davidson TB, et al. Mutations in the sonic hedgehog pathway cause macrocephaly-associated conditions due to crosstalk to the PI3K/AKT/mTOR pathway. Am J Med Genet A. 2019;179(12):2517-31. Epub 20191022. doi: 10.1002/ajmg.a.61368. PubMed PMID: 31639285; PubMed Central PMCID: 7346528.

154. Balk K, Biesecker LG. The clinical atlas of Greig cephalopolysyndactyly syndrome. Am J Med Genet A. 2008;146A(5):548-57. doi: 10.1002/ajmg.a.32167. PubMed PMID: 18241058.

155. Dafinger C, Liebau MC, Elsayed SM, Hellenbroich Y, Boltshauser E, Korenke GC, et al. Mutations in KIF7 link Joubert syndrome with Sonic Hedgehog signaling and microtubule dynamics. J Clin Invest. 2011;121(7):2662-7. doi: 10.1172/JCI43639. PubMed PMID: 21633164; PubMed Central PMCID: 3223820.

156. Johnson RL, Rothman AL, Xie J, Goodrich LV, Bare JW, Bonifas JM, et al. Human homolog of patched, a candidate gene for the basal cell nevus syndrome. Science. 1996;272(5268):1668-71. doi: 10.1126/science.272.5268.1668. PubMed PMID: 8658145.

157. GORLIN RJ, GOLTZ RW. Multiple nevoid basal-cell epithelioma, jaw cysts and bifid rib. A syndrome. N Engl J Med. 1960;262:908-12. doi: 10.1056/NEJM196005052621803. PubMed PMID: 13851319.

158. Cottereau E, Mortemousque I, Moizard MP, Bürglen L, Lacombe D, Gilbert-Dussardier B, et al. Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature. Am J Med Genet C Semin Med Genet. 2013;163C(2):92-105. Epub 20130418. doi: 10.1002/ajmg.c.31360. PubMed PMID: 23606591.

159. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. Nat Genet. 2008;40(12):1466-71. doi: 10.1038/ng.279. PubMed PMID: 19029900; PubMed Central PMCID: 2680128.

160. Tatton-Brown K, Murray A, Hanks S, Douglas J, Armstrong R, Banka S, et al. Weaver syndrome and EZH2 mutations: Clarifying the clinical phenotype. Am J Med Genet A. 2013;161A(12):2972-80. Epub 20131108. doi: 10.1002/ajmg.a.36229. PubMed PMID: 24214728.

161. Tatton-Brown K, Zachariou A, Loveday C, Renwick A, Mahamdallie S, Aksglaede L, et al. The Tatton-Brown-Rahman Syndrome: A clinical study of 55 individuals with. Wellcome Open Res. 2018;3:46. Epub 20180423. doi: 10.12688/wellcomeopenres.14430.1. PubMed PMID: 29900417; PubMed Central PMCID: 5964628.

162. Klaassens M, Morrogh D, Rosser EM, Jaffer F, Vreeburg M, Bok LA, et al. Malan syndrome: Sotos-like overgrowth with de novo NFIX sequence variants and deletions in six new patients and a review of the literature. Eur J Hum Genet.

2015;23(5):610-5. Epub 20140813. doi: 10.1038/ejhg.2014.162. PubMed PMID: 25118028; PubMed Central PMCID: 4402637.

163.   Budny B, Chen W, Omran H, Fliegauf M, Tzschach A, Wisniewska M, et al. A novel X-linked recessive mental retardation syndrome comprising macrocephaly and ciliary dysfunction is allelic to oral-facial-digital type I syndrome. Hum Genet. 2006;120(2):171-8. Epub 20060617. doi: 10.1007/s00439-006-0210-5. PubMed PMID: 16783569.

164.   Patton MA, Afzal AR. Robinow syndrome. J Med Genet. 2002;39(5):305-10. doi: 10.1136/jmg.39.5.305. PubMed PMID: 12011143; PubMed Central PMCID: 1735132.

165.   Bellus GA, Hefferon TW, Ortiz de Luna RI, Hecht JT, Horton WA, Machado M, et al. Achondroplasia is defined by recurrent G380R mutations of FGFR3. Am J Hum Genet. 1995;56(2):368-73. PubMed PMID: 7847369; PubMed Central PMCID: 1801129.

166.   Jansen LA, Mirzaa GM, Ishak GE, O'Roak BJ, Hiatt JB, Roden WH, et al. PI3K/AKT pathway mutations cause a spectrum of brain malformations from megalencephaly to focal cortical dysplasia. Brain. 2015;138(Pt 6):1613-28. Epub 20150225. doi: 10.1093/brain/awv045. PubMed PMID: 25722288; PubMed Central PMCID: 4614119.

167.   Orlova KA, Parker WE, Heuer GG, Tsai V, Yoon J, Baybis M, et al. STRADalpha deficiency results in aberrant mTORC1 signaling during corticogenesis in humans and mice. J Clin Invest. 2010;120(5):1591-602. Epub 2010/04/26. doi: 10.1172/JCI41592. PubMed PMID: 20424326; PubMed Central PMCID: 2860905.

168.   Pavone P, Praticò AD, Rizzo R, Corsello G, Ruggieri M, Parano E, et al. A clinical review on megalencephaly: A large brain as a possible sign of cerebral impairment. Medicine (Baltimore). 2017;96(26):e6814. doi: 10.1097/MD.0000000000006814. PubMed PMID: 28658095; PubMed Central PMCID: 5500017.

169.   Hill AD, Chang BS, Hill RS, Garraway LA, Bodell A, Sellers WR, et al. A 2-Mb critical region implicated in the microcephaly associated with terminal 1q deletion syndrome. Am J Med Genet A. 2007;143A(15):1692-8. doi: 10.1002/ajmg.a.31776. PubMed PMID: 17603806.

170.   Pang T, Atefy R, Sheen V. Malformations of cortical development. Neurologist. 2008;14(3):181-91. doi: 10.1097/NRL.0b013e31816606b9. PubMed PMID: 18469675; PubMed Central PMCID: 3547618.

171.   Fry AE, Cushion TD, Pilz DT. The genetics of lissencephaly. Am J Med Genet C Semin Med Genet. 2014;166C(2):198-210. Epub 20140523. doi: 10.1002/ajmg.c.31402. PubMed PMID: 24862549.

172.   Chakraborti S, Natarajan K, Curiel J, Janke C, Liu J. The emerging role of the tubulin code: From the tubulin molecule to neuronal function and disease.

Cytoskeleton (Hoboken). 2016;73(10):521-50. Epub 20160509. doi: 10.1002/cm.21290. PubMed PMID: 26934450.

173. Bahi-Buisson N, Poirier K, Boddaert N, Saillour Y, Castelnau L, Philip N, et al. Refinement of cortical dysgeneses spectrum associated with TUBA1A mutations. J Med Genet. 2008;45(10):647-53. Epub 20080826. doi: 10.1136/jmg.2008.058073. PubMed PMID: 18728072.

174. Crino PB. Focal Cortical Dysplasia. Semin Neurol. 2015;35(3):201-8. Epub 20150610. doi: 10.1055/s-0035-1552617. PubMed PMID: 26060899; PubMed Central PMCID: 6413691.

175. Stutterd CA, Leventer RJ. Polymicrogyria: a common and heterogeneous malformation of cortical development. Am J Med Genet C Semin Med Genet. 2014;166C(2):227-39. Epub 20140528. doi: 10.1002/ajmg.c.31399. PubMed PMID: 24888723.

176. Judkins AR, Martinez D, Ferreira P, Dobyns WB, Golden JA. Polymicrogyria includes fusion of the molecular layer and decreased neuronal populations but normal cortical laminar organization. J Neuropathol Exp Neurol. 2011;70(6):438-43. doi: 10.1097/NEN.0b013e31821ccf1c. PubMed PMID: 21572338; PubMed Central PMCID: 3113653.

177. Hofman J, Hutny M, Sztuba K, Paprocka J. Corpus Callosum Agenesis: An Insight into the Etiology and Spectrum of Symptoms. Brain Sci. 2020;10(9). Epub 20200909. doi: 10.3390/brainsci10090625. PubMed PMID: 32916978; PubMed Central PMCID: 7565833.

178. Paul LK, Brown WS, Adolphs R, Tyszka JM, Richards LJ, Mukherjee P, et al. Agenesis of the corpus callosum: genetic, developmental and functional aspects of connectivity. Nat Rev Neurosci. 2007;8(4):287-99. doi: 10.1038/nrn2107. PubMed PMID: 17375041.

179. Pilz DT, Matsumoto N, Minnerath S, Mills P, Gleeson JG, Allen KM, et al. LIS1 and XLIS (DCX) mutations cause most classical lissencephaly, but different patterns of malformation. Hum Mol Genet. 1998;7(13):2029-37. doi: 10.1093/hmg/7.13.2029. PubMed PMID: 9817918.

180. Lee GH, D'Arcangelo G. New Insights into Reelin-Mediated Signaling Pathways. Front Cell Neurosci. 2016;10:122. Epub 20160509. doi: 10.3389/fncel.2016.00122. PubMed PMID: 27242434; PubMed Central PMCID: 4860420.

181. Sheen VL, Jansen A, Chen MH, Parrini E, Morgan T, Ravenscroft R, et al. Filamin A mutations cause periventricular heterotopia with Ehlers-Danlos syndrome. Neurology. 2005;64(2):254-62. doi: 10.1212/01.WNL.0000149512.79621.DF. PubMed PMID: 15668422.

182. Poirier K, Lebrun N, Broix L, Tian G, Saillour Y, Boscheron C, et al. Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. Nat Genet. 2013;45(6):639-47. Epub

20130421. doi: 10.1038/ng.2613. PubMed PMID: 23603762; PubMed Central PMCID: 3826256.

183.   Heinzen EL, O'Neill AC, Zhu X, Allen AS, Bahlo M, Chelly J, et al. De novo and inherited private variants in MAP1B in periventricular nodular heterotopia. PLoS Genet. 2018;14(5):e1007281. Epub 20180508. doi: 10.1371/journal.pgen.1007281. PubMed PMID: 29738522; PubMed Central PMCID: 5965900.

184.   Cappello S, Gray MJ, Badouel C, Lange S, Einsiedler M, Srour M, et al. Mutations in genes encoding the cadherin receptor-ligand pair DCHS1 and FAT4 disrupt cerebral cortical development. Nat Genet. 2013;45(11):1300-8. Epub 20130922. doi: 10.1038/ng.2765. PubMed PMID: 24056717.

185.   Smogavec M, Cleall A, Hoyer J, Lederer D, Nassogne MC, Palmer EE, et al. Eight further individuals with intellectual disability and epilepsy carrying bi-allelic CNTNAP2 aberrations allow delineation of the mutational and phenotypic spectrum. J Med Genet. 2016;53(12):820-7. Epub 2016/07/20. doi: 10.1136/jmedgenet-2016-103880. PubMed PMID: 27439707.

186.   Weckhuysen S, Marsan E, Lambrecq V, Marchal C, Morin-Brureau M, An-Gourfinkel I, et al. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. Epilepsia. 2016;57(6):994-1003. Epub 2016/05/13. doi: 10.1111/epi.13391. PubMed PMID: 27173016.

187.   Yu TW, Mochida GH, Tischfield DJ, Sgaier SK, Flores-Sarnat L, Sergi CM, et al. Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. Nat Genet. 2010;42(11):1015-20. Epub 2010/10/03. doi: 10.1038/ng.683. PubMed PMID: 20890278; PubMed Central PMCID: 2969850.

188.   Baulac S, Lenk GM, Dufresnois B, Ouled Amar Bencheikh B, Couarch P, Renard J, et al. Role of the phosphoinositide phosphatase FIG4 gene in familial epilepsy with polymicrogyria. Neurology. 2014;82(12):1068-75. Epub 20140305. doi: 10.1212/WNL.0000000000000241. PubMed PMID: 24598713; PubMed Central PMCID: 3962989.

189.   Aldinger KA, Doherty D. The genetics of cerebellar malformations. Semin Fetal Neonatal Med. 2016;21(5):321-32. Epub 20160507. doi: 10.1016/j.siny.2016.04.008. PubMed PMID: 27160001; PubMed Central PMCID: 5035570.

190.   Hong SE, Shugart YY, Huang DT, Shahwan SA, Grant PE, Hourihane JO, et al. Autosomal recessive lissencephaly with cerebellar hypoplasia is associated with human RELN mutations. Nat Genet. 2000;26(1):93-6. doi: 10.1038/79246. PubMed PMID: 10973257.

191.   Poretti A, Boltshauser E. Terminology in morphological anomalies of the cerebellum does matter. Cerebellum Ataxias. 2015;2:8. Epub 20150707. doi: 10.1186/s40673-015-0027-x. PubMed PMID: 26331051; PubMed Central PMCID: 4552363.

192.   Rudnik-Schöneborn S, Barth PG, Zerres K. Pontocerebellar hypoplasia. Am J Med Genet C Semin Med Genet. 2014;166C(2):173-83. Epub 20140612. doi: 10.1002/ajmg.c.31403. PubMed PMID: 24924738.

193.   WHO. The International Statistical Classification of Diseases and Related Health Problems, 11th edition, World Health Organisation   2019.

194.   APA. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. 2013.

195.   Morris-Rosendahl DJ, Crocq MA. Neurodevelopmental disorders-the history and future of a diagnostic concept. Dialogues Clin Neurosci. 2020;22(1):65-72. doi: 10.31887/DCNS.2020.22.1/macrocq. PubMed PMID: 32699506; PubMed Central PMCID: 7365295.

196.   WHO. Definition: intellectual disability: World Health Organisation: regional office for Europe; 2022. Available from: https://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/news/news/2010/15/childrens-right-to-family-life/definition-intellectual-disability.

197.   Roid G, Pomplun M. Stanford Binet intelligence scales - Fifth edition. Itasca, Illinois: Riverside Publishing; 2012.

198.   Wechsler D. *Wechsler intelligence scale for children – Fourth edition (WISC-IV).* San Antonio, Texas: American Psychological Association; 2003.

199.   Wechsler D. *Wechsler Adult Intelligence Scale - Fourth edition (WAIS-IV).* San Antonio, Texas: Pearson Assessment; 2008.

200.   Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. Res Dev Disabil. 2011;32(2):419-36. Epub 20110113. doi: 10.1016/j.ridd.2010.12.018. PubMed PMID: 21236634.

201.   Salvador-Carulla L, Symonds S. Health services use and costs in people with intellectual disability: building a context knowledge base for evidence-informed policy. Curr Opin Psychiatry. 2016;29(2):89-94. doi: 10.1097/YCO.0000000000000237. PubMed PMID: 26808613.

202.   Reichenberg A, Cederlöf M, McMillan A, Trzaskowski M, Kapra O, Fruchter E, et al. Discontinuity in the genetic and environmental causes of the intellectual disability spectrum. Proc Natl Acad Sci U S A. 2016;113(4):1098-103. Epub 20151228. doi: 10.1073/pnas.1508093112. PubMed PMID: 26711998; PubMed Central PMCID: 4743770.

203.   Majnemer A, Shevell MI. Diagnostic yield of the neurologic assessment of the developmentally delayed child. J Pediatr. 1995;127(2):193-9. doi: 10.1016/s0022-3476(95)70294-6. PubMed PMID: 7543566.

204.   Shevell MI, Majnemer A, Rosenbaum P, Abrahamowicz M. Etiologic yield of subspecialists' evaluation of young children with global developmental delay. J Pediatr. 2000;136(5):593-8. doi: 10.1067/mpd.2000.104817. PubMed PMID: 10802489.

205.   Südhof TC. Neuroligins and neurexins link synaptic function to cognitive disease. Nature. 2008;455(7215):903-11. doi: 10.1038/nature07456. PubMed PMID: 18923512; PubMed Central PMCID: 2673233.

206.   Cristino AS, Williams SM, Hawi Z, An JY, Bellgrove MA, Schwartz CE, et al. Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. Mol Psychiatry. 2014;19(3):294-301. Epub 20130226. doi: 10.1038/mp.2013.16. PubMed PMID: 23439483.

207.   Mithyantha R, Kneen R, McCann E, Gladstone M. Current evidence-based recommendations on investigating children with global developmental delay. Arch Dis Child. 2017;102(11):1071-6. doi: 10.1136/archdischild-2016-311271. PubMed PMID: 29054862; PubMed Central PMCID: 5738593.

208.   Rodriguez DP, Poussaint TY. Neuroimaging of the child with developmental delay. Top Magn Reson Imaging. 2007;18(1):75-92. doi: 10.1097/RMR.0b013e3180f63511. PubMed PMID: 17607144.

209.   Barkovich AJ. Imaging of the Newborn Brain. Semin Pediatr Neurol. 2019;32:100766. Epub 20190906. doi: 10.1016/j.spen.2019.08.002. PubMed PMID: 31813522.

210.   Ten Kate LP, Al-Gazali L, Anand S, Bittles A, Cassiman JJ, Christianson A, et al. Community genetics. Its definition 2010. J Community Genet. 2010;1(1):19-22. Epub 20100331. doi: 10.1007/s12687-010-0007-z. PubMed PMID: 21475671; PubMed Central PMCID: 3063846.

211.   Nolt S. A History of the Amish. 3rd ed2016.

212.   College E. Amish Population Profile 2021: Amish Studies: The Young Center for Anabaptist and Pietist Studies; 2021 [25/01/2022]. Available from: https://groups.etown.edu/amishstudies/statistics/amish-population-profile-2021/.

213.   Kraybill D, Johnson-Weiner K, Nolt S. The Amish: Johns Hopkins University Press; 2013.

214.   Braman S. Map of Amish Counties 2020 2021. Available from: https://bramanswanderings.com/2021/01/23/map-of-amish-counties-2020/.

215.   Alkuraya FS. The application of next-generation sequencing in the autozygosity mapping of human recessive diseases. Hum Genet. 2013;132(11):1197-211. Epub 20130802. doi: 10.1007/s00439-013-1344-x. PubMed PMID: 23907654.

216.   Lander ES, Botstein D. Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science. 1987;236(4808):1567-70. doi: 10.1126/science.2884728. PubMed PMID: 2884728.

217. Cooper GM, Mefford HC. Detection of copy number variation using SNP genotyping. Methods Mol Biol. 2011;767:243-52. doi: 10.1007/978-1-61779-201-4_18. PubMed PMID: 21822880.

218. Zollo M, Ahmed M, Ferrucci V, Salpietro V, Asadzadeh F, Carotenuto M, et al. PRUNE is crucial for normal brain development and mutated in microcephaly with neurodevelopmental impairment. Brain. 2017;140(4):940-52. doi: 10.1093/brain/awx014. PubMed PMID: 28334956; PubMed Central PMCID: 5382943.

219. Ahmed MY, Al-Khayat A, Al-Murshedi F, Al-Futaisi A, Chioza BA, Pedro Fernandez-Murray J, et al. A mutation of EPT1 (SELENOI) underlies a new disorder of Kennedy pathway phospholipid biosynthesis. Brain. 2017;140(3):547-54. doi: 10.1093/brain/aww318. PubMed PMID: 28052917; PubMed Central PMCID: 5382949.

220. MCKUSICK VA, HOSTETLER JA, EGELAND JA. GENETIC STUDIES OF THE AMISH, BACKGROUND AND POTENTIALITIES. Bull Johns Hopkins Hosp. 1964;115:203-22. PubMed PMID: 14209042.

221. Francomano CA, McKusick VA, Biesecker LG. Medical genetic studies in the Amish: historical perspective. Am J Med Genet C Semin Med Genet. 2003;121C(1):1-4. doi: 10.1002/ajmg.c.20001. PubMed PMID: 12888981.

222. McKusick V. Medical genetic studies of the Amish: selected papers: Johns Hopkins University Press; 1978.

223. Ammous Z, Rawlins LE, Jones H, Leslie JS, Wenger O, Scott E, et al. A biallelic SNIP1 Amish founder variant causes a recognizable neurodevelopmental disorder. PLoS Genet. 2021;17(9):e1009803. Epub 20210927. doi: 10.1371/journal.pgen.1009803. PubMed PMID: 34570759; PubMed Central PMCID: 8496849.

224. Leslie JS, Rawlins LE, Chioza BA, Olubodun OR, Salter CG, Fasham J, et al. MNS1 variant associated with situs inversus and male infertility. Eur J Hum Genet. 2020;28(1):50-5. Epub 20190918. doi: 10.1038/s41431-019-0489-z. PubMed PMID: 31534215; PubMed Central PMCID: 6906318.

225. Harlalka GV, Baple EL, Cross H, Kühnle S, Cubillos-Rojas M, Matentzoglu K, et al. Mutation of HERC2 causes developmental delay with Angelman-like features. J Med Genet. 2013;50(2):65-73. Epub 2012/12/14. doi: 10.1136/jmedgenet-2012-101367. PubMed PMID: 23243086.

226. UN UNPF. World Population Dashboard: Pakistan 2021 [27th January 2022]. Available from: https://www.unfpa.org/data/world-population/PK.

227. Bureau' USC. World Population Clock: Most populous countries 2021 [27th January 2022]. Available from: https://www.census.gov/popclock/world.

228. Sathar Z, Royan R, Bongaarts J. Capturing the demographic dividend in Pakistan: Pakistan Population Council; 2013.

229.   Hussain R. The effect of religious, cultural and social identity on population genetic structure among Muslims in Pakistan. Ann Hum Biol. 2005;32(2):145-53. doi: 10.1080/03014460500075167. PubMed PMID: 16096210.

230.   PBS. Pakistan Bureau of Statistics: Final Results of Census 2017: Government of Pakistan; 2017. Available from: https://www.pbs.gov.pk/content/final-results-census-2017-0.

231.   Mehdil S, Qamarl R, Ayub l Q, Khaliql S, Mansoorl AI, smaill M, et al. THE ORIGINS OF PAKISTANI POPULATIONS

Evidence from Y Chromosome Markers. In: Papiha D, and Chakraborty, editor. Genomic Diversity: Applications in Human Population Genetics. New York: Kluwer Academic / Plenum Publishers; 1999.

232.   Bittles AH, Small NA. CONSANGUINITY, GENETICS AND DEFINITIONS OF KINSHIP IN THE UK PAKISTANI POPULATION. J Biosoc Sci. 2016;48(6):844-54. Epub 20151228. doi: 10.1017/S0021932015000449. PubMed PMID: 26707179.

233.   Iqbal Z, van Bokhoven H. Identifying genes responsible for intellectual disability in consanguineous families. Hum Hered. 2014;77(1-4):150-60. Epub 20140729. doi: 10.1159/000360539. PubMed PMID: 25060278.

234.   Jabeen N, Malik S. Consanguinity and its sociodemographic differentials in Bhimber District, Azad Jammu and Kashmir, Pakistan. J Health Popul Nutr. 2014;32(2):301-13. PubMed PMID: 25076667; PubMed Central PMCID: 4216966.

235.   Riaz M, Tiller J, Ajmal M, Azam M, Qamar R, Lacaze P. Implementation of public health genomics in Pakistan. Eur J Hum Genet. 2019;27(10):1485-92. Epub 20190517. doi: 10.1038/s41431-019-0428-z. PubMed PMID: 31101884; PubMed Central PMCID: 6777461.

236.   Gustavson KH. Prevalence and aetiology of congenital birth defects, infant mortality and mental retardation in Lahore, Pakistan: a prospective cohort study. Acta Paediatr. 2005;94(6):769-74. doi: 10.1111/j.1651-2227.2005.tb01981.x. PubMed PMID: 16188785.

237.   Wang W, Sullivan SG, Ahmed S, Chandler D, Zhivotovsky LA, Bittles AH. A genome-based study of consanguinity in three co-resident endogamous Pakistan communities. Ann Hum Genet. 2000;64(Pt 1):41-9. doi: 10.1017/S0003480000007946. PubMed PMID: 11246460.

238.   Khan S, Lin S, Harlalka GV, Ullah A, Shah K, Khalid S, et al. BBS5 and INPP5E mutations associated with ciliopathy disorders in families from Pakistan. Ann Hum Genet. 2019;83(6):477-82. Epub 20190607. doi: 10.1111/ahg.12336. PubMed PMID: 31173343.

239.   Qasim I, Ahmad B, Khan MA, Khan N, Muhammad N, Basit S, et al. Pakistan Genetic Mutation Database (PGMD); A centralized Pakistani mutome

data source. Eur J Med Genet. 2018;61(4):204-8. Epub 20171207. doi: 10.1016/j.ejmg.2017.11.015. PubMed PMID: 29223505.

240. Kochinke K, Zweier C, Nijhof B, Fenckova M, Cizek P, Honti F, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. Am J Hum Genet. 2016;98(1):149-64. doi: 10.1016/j.ajhg.2015.11.024. PubMed PMID: 26748517; PubMed Central PMCID: 4716705.

241. Anazi S, Maddirevula S, Faqeih E, Alsedairy H, Alzahrani F, Shamseldin HE, et al. Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. Mol Psychiatry. 2017;22(4):615-24. Epub 20160719. doi: 10.1038/mp.2016.113. PubMed PMID: 27431290.

242. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. Lancet. 2015;385(9975):1305-14. Epub 2014/12/17. doi: 10.1016/S0140-6736(14)61705-0. PubMed PMID: 25529582; PubMed Central PMCID: 4392068.

243. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature. 2014;511(7509):344-7. Epub 20140604. doi: 10.1038/nature13394. PubMed PMID: 24896178.

244. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36(10):928-30. Epub 2015/08/13. doi: 10.1002/humu.22844. PubMed PMID: 26220891; PubMed Central PMCID: 4833888.

245. Pan H, Cole T. LMS growth, a Microsoft Excel add-in to access growth references based on the LMS method. Version 2.77. 2012.

246. Korkman M, Kirk U, Kemp S. *Developmental Neuropsychological Assessment-Second edition (NEPSY-II).* San Antonio, Texas: Harcourt Assessment; 2007.

247. Randolph C. Repeatable Battery for the Assessment of Neuropsychological Status (RBANS)- Update. Bloomington, Minnesota: Pearson Education; 2012.

248. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006;7:3. Epub 20060131. doi: 10.1186/1471-2199-7-3. PubMed PMID: 16448564; PubMed Central PMCID: 1413964.

249. Korbie DJ, Mattick JS. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat Protoc. 2008;3(9):1452-6. doi: 10.1038/nprot.2008.133. PubMed PMID: 18772872.

250. Hardjasa A, Ling M, Ma K, Yu H. Investigating the effects of DMSO on PCR fidelity using a restriction digest-based method. J Exp Microbiol Immunol. 2010;14:161-4.

251. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv. 2013;13033997.

252. BI. Broad Institute. Picard Toolkit v2.15.0 2021. Available from: http://broadinstitute.github.io/picard/.

253. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-303. Epub 20100719. doi: 10.1101/gr.107524.110. PubMed PMID: 20644199; PubMed Central PMCID: 2928508.

254. Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, et al. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics. 2012;28(21):2747-54. Epub 20120831. doi: 10.1093/bioinformatics/bts526. PubMed PMID: 22942019; PubMed Central PMCID: 3476336.

255. Laver T, De Franco E, Johnson M, Patel K, Ellard S, Weedon M, et al. SavvyCNV: genome-wide CNV calling from off-target reads. bioRxiv preprint. 2019;617605.

256. Whiffin N, Karczewski KJ, Zhang X, Chothani S, Smith MJ, Evans DG, et al. Characterising the loss-of-function impact of 5' untranslated region variants in 15,708 individuals. Nat Commun. 2020;11(1):2523. Epub 20200527. doi: 10.1038/s41467-019-10717-9. PubMed PMID: 32461616; PubMed Central PMCID: 7253449.

257. Vincze T, Posfai J, Roberts RJ. NEBcutter: A program to cleave DNA with restriction enzymes. Nucleic Acids Res. 2003;31(13):3688-91. doi: 10.1093/nar/gkg526. PubMed PMID: 12824395; PubMed Central PMCID: 168933.

258. Sacher M, Barrowman J, Wang W, Horecka J, Zhang Y, Pypaert M, et al. TRAPP I implicated in the specificity of tethering in ER-to-Golgi transport. Mol Cell. 2001;7(2):433-42. PubMed PMID: 11239471.

259. Cai H, Yu S, Menon S, Cai Y, Lazarova D, Fu C, et al. TRAPPI tethers COPII vesicles by binding the coat subunit Sec23. Nature. 2007;445(7130):941-4. Epub 2007/02/07. doi: 10.1038/nature05527. PubMed PMID: 17287728.

260. Lynch-Day MA, Bhandari D, Menon S, Huang J, Cai H, Bartholomew CR, et al. Trs85 directs a Ypt1 GEF, TRAPPIII, to the phagophore to promote autophagy. Proc Natl Acad Sci U S A. 2010;107(17):7811-6. Epub 2010/04/07. doi: 10.1073/pnas.1000063107. PubMed PMID: 20375281; PubMed Central PMCID: 2867920.

261.	Brunet S, Sacher M. In sickness and in health: the role of TRAPP and associated proteins in disease. Traffic. 2014;15(8):803-18. Epub 20140701. doi: 10.1111/tra.12183. PubMed PMID: 24917561.

262.	Cuenca A, Insinna C, Zhao H, John P, Weiss MA, Lu Q, et al. The C7orf43/TRAPPC14 component links the TRAPPII complex to Rabin8 for preciliary vesicle tethering at the mother centriole during ciliogenesis. J Biol Chem.	2019;294(42):15418-34.	Epub	2019/08/29.	doi: 10.1074/jbc.RA119.008615. PubMed PMID: 31467083; PubMed Central PMCID: 6802515.

263.	Riedel F, Galindo A, Muschalik N, Munro S. The two TRAPP complexes of metazoans have distinct roles and act on different Rab GTPases. J Cell Biol. 2018;217(2):601-17. Epub 2017/12/22. doi: 10.1083/jcb.201705068. PubMed PMID: 29273580; PubMed Central PMCID: 5800803.

264.	Jenkins ML, Harris NJ, Dalwadi U, Fleming KD, Ziemianowicz DS, Rafiei A, et al. The substrate specificity of the human TRAPPII complex's Rab-guanine nucleotide exchange factor activity. Commun Biol. 2020;3(1):735. Epub 2020/12/04. doi: 10.1038/s42003-020-01459-2. PubMed PMID: 33277614; PubMed Central PMCID: 7719173.

265.	Scrivens PJ, Noueihed B, Shahrzad N, Hul S, Brunet S, Sacher M. C4orf41 and TTC-15 are mammalian TRAPP components with a role at an early stage in ER-to-Golgi trafficking. Mol Biol Cell. 2011;22(12):2083-93. Epub 2011/04/27. doi: 10.1091/mbc.E10-11-0873. PubMed PMID: 21525244; PubMed Central PMCID: 3113772.

266.	Yamasaki A, Menon S, Yu S, Barrowman J, Meerloo T, Oorschot V, et al. mTrs130 is a component of a mammalian TRAPPII complex, a Rab1 GEF that binds to COPI-coated vesicles. Mol Biol Cell. 2009;20(19):4205-15. Epub 2009/08/05. doi: 10.1091/mbc.e09-05-0387. PubMed PMID: 19656848; PubMed Central PMCID: 2754934.

267.	Li C, Luo X, Zhao S, Siu GK, Liang Y, Chan HC, et al. COPI-TRAPPII activates Rab18 and regulates its lipid droplet association. EMBO J. 2017;36(4):441-57. Epub 2016/12/21. doi: 10.15252/embj.201694866. PubMed PMID: 28003315; PubMed Central PMCID: 5694949.

268.	Zhao S, Li CM, Luo XM, Siu GK, Gan WJ, Zhang L, et al. Mammalian TRAPPIII Complex positively modulates the recruitment of Sec13/31 onto COPII vesicles. Sci Rep. 2017;7:43207. Epub 2017/02/27. doi: 10.1038/srep43207. PubMed PMID: 28240221; PubMed Central PMCID: 5327430.

269.	Westlake CJ, Baye LM, Nachury MV, Wright KJ, Ervin KE, Phu L, et al. Primary cilia membrane assembly is initiated by Rab11 and transport protein particle II (TRAPPII) complex-dependent trafficking of Rabin8 to the centrosome. Proc Natl Acad Sci U S A. 2011;108(7):2759-64. Epub 2011/01/27. doi: 10.1073/pnas.1018823108. PubMed PMID: 21273506; PubMed Central PMCID: 3041065.

270.   Zhang C, Li C, Siu GKY, Luo X, Yu S. Distinct Roles of TRAPPC8 and TRAPPC12 in Ciliogenesis via Their Interactions With OFD1. Front Cell Dev Biol. 2020;8:148. Epub 20200317. doi: 10.3389/fcell.2020.00148. PubMed PMID: 32258032; PubMed Central PMCID: 7090148.

271.   Ghosh AK, Majumder M, Steele R, White RA, Ray RB. A novel 16-kilodalton cellular protein physically interacts with and antagonizes the functional activity of c-myc promoter-binding protein 1. Mol Cell Biol. 2001;21(2):655-62. doi: 10.1128/MCB.21.2.655-662.2001. PubMed PMID: 11134351; PubMed Central PMCID: 86643.

272.   DeRossi C, Vacaru A, Rafiq R, Cinaroglu A, Imrie D, Nayar S, et al. trappc11 is required for protein glycosylation in zebrafish and humans. Mol Biol Cell. 2016;27(8):1220-34. Epub 20160224. doi: 10.1091/mbc.E15-08-0557. PubMed PMID: 26912795; PubMed Central PMCID: 4831877.

273.   Larson AA, Baker PR, Milev MP, Press CA, Sokol RJ, Cox MO, et al. TRAPPC11 and GOSR2 mutations associate with hypoglycosylation of α-dystroglycan and muscular dystrophy. Skelet Muscle. 2018;8(1):17. Epub 2018/05/31. doi: 10.1186/s13395-018-0163-0. PubMed PMID: 29855340; PubMed Central PMCID: 5984345.

274.   Zong M, Satoh A, Yu MK, Siu KY, Ng WY, Chan HC, et al. TRAPPC9 mediates the interaction between p150 and COPII vesicles at the target membrane. PLoS One. 2012;7(1):e29995. Epub 2012/01/18. doi: 10.1371/journal.pone.0029995. PubMed PMID: 22279557; PubMed Central PMCID: 3261171.

275.   Hu WH, Pendergast JS, Mo XM, Brambilla R, Bracchi-Ricard V, Li F, et al. NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation. J Biol Chem. 2005;280(32):29233-41. Epub 2005/06/10. doi: 10.1074/jbc.M501670200. PubMed PMID: 15951441; PubMed Central PMCID: 3707486.

276.   Milev MP, Graziano C, Karall D, Kuper WFE, Al-Deri N, Cordelli DM, et al. Bi-allelic mutations in TRAPPC2L result in a neurodevelopmental disorder and have an impact on Rab11 in fibroblasts. J Med Genet. 2018;55(11):753-64. Epub 2018/08/17. doi: 10.1136/jmedgenet-2018-105441. PubMed PMID: 30120216.

277.   Fiedler J, Le Merrer M, Mortier G, Heuertz S, Faivre L, Brenner RE. X-linked spondyloepiphyseal dysplasia tarda: Novel and recurrent mutations in 13 European families. Hum Mutat. 2004;24(1):103. doi: 10.1002/humu.9254. PubMed PMID: 15221797.

278.   Venditti R, Scanu T, Santoro M, Di Tullio G, Spaar A, Gaibisso R, et al. Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. Science. 2012;337(6102):1668-72. doi: 10.1126/science.1224947. PubMed PMID: 23019651; PubMed Central PMCID: 3471527.

279.   Gedeon AK, Tiller GE, Le Merrer M, Heuertz S, Tranebjaerg L, Chitayat D, et al. The molecular basis of X-linked spondyloepiphyseal dysplasia tarda. Am

J Hum Genet. 2001;68(6):1386-97. Epub 2001/05/08. doi: 10.1086/320592. PubMed PMID: 11349230; PubMed Central PMCID: 1226125.

280. Al-Deri N, Okur V, Ahimaz P, Milev M, Valivullah Z, Hagen J, et al. A novel homozygous variant in *TRAPPC2L* results in a neurodevelopmental disorder and disrupts TRAPP complex function. J Med Genet. 2020. Epub 2020/08/25. doi: 10.1136/jmedgenet-2020-107016. PubMed PMID: 32843486.

281. Shaheen R, Szymanska K, Basu B, Patel N, Ewida N, Faqeih E, et al. Characterizing the morbid genome of ciliopathies. Genome Biol. 2016;17(1):242. Epub 20161128. doi: 10.1186/s13059-016-1099-5. PubMed PMID: 27894351; PubMed Central PMCID: 5126998.

282. Van Bergen NJ, Guo Y, Al-Deri N, Lipatova Z, Stanga D, Zhao S, et al. Deficiencies in vesicular transport mediated by TRAPPC4 are associated with severe syndromic intellectual disability. Brain. 2020;143(1):112-30. doi: 10.1093/brain/awz374. PubMed PMID: 31794024; PubMed Central PMCID: 6935753.

283. Ghosh SG, Scala M, Beetz C, Helman G, Stanley V, Yang X, et al. A relatively common homozygous TRAPPC4 splicing variant is associated with an early-infantile neurodegenerative syndrome. Eur J Hum Genet. 2021;29(2):271-9. Epub 20200908. doi: 10.1038/s41431-020-00717-5. PubMed PMID: 32901138; PubMed Central PMCID: 7868361.

284. Saad AK, Marafi D, Mitani T, Jolly A, Du H, Elbendary HM, et al. Biallelic in-frame deletion in TRAPPC4 in a family with developmental delay and cerebellar atrophy. Brain. 2020;143(10):e83. doi: 10.1093/brain/awaa256. PubMed PMID: 33011761; PubMed Central PMCID: 7586085.

285. Majethia P, Do Rosario MC, Kaur P, Karanvir, Shankar R, Sharma S, et al. Further evidence of muscle involvement in neurodevelopmental disorder with epilepsy, spasticity, and brain atrophy. Ann Hum Genet. 2022;86(2):94-101. Epub 20211208. doi: 10.1111/ahg.12452. PubMed PMID: 34878169.

286. Kaur P, Kadavigere R, Girisha KM, Shukla A. Recurrent bi-allelic splicing variant c.454+3A>G in TRAPPC4 is associated with progressive encephalopathy and muscle involvement. Brain. 2020;143(4):e29. doi: 10.1093/brain/awaa046. PubMed PMID: 32125366.

287. Mohamoud HS, Ahmed S, Jelani M, Alrayes N, Childs K, Vadgama N, et al. A missense mutation in TRAPPC6A leads to build-up of the protein, in patients with a neurodevelopmental syndrome and dysmorphic features. Sci Rep. 2018;8(1):2053. Epub 2018/02/01. doi: 10.1038/s41598-018-20658-w. PubMed PMID: 29391579; PubMed Central PMCID: 5794855.

288. Marin-Valencia I, Novarino G, Johansen A, Rosti B, Issa MY, Musaev D, et al. A homozygous founder mutation in TRAPPC6B associates with a neurodevelopmental disorder characterised by microcephaly, epilepsy and autistic features. J Med Genet. 2018;55(1):48-54. Epub 2017/06/16. doi: 10.1136/jmedgenet-2017-104627. PubMed PMID: 28626029; PubMed Central PMCID: 6056005.

289. Nair P, El-Bazzal L, Mansour H, Sabbagh S, Al-Ali MT, Gambarini A, et al. Further Delineation of the *TRAPPC6B* Disorder: Report of a New Family and Review. J Pediatr Genet. 2019;8(4):252-6. Epub 20190730. doi: 10.1055/s-0039-1693664. PubMed PMID: 31687267; PubMed Central PMCID: 6824894.

290. Rasool IG, Zahoor MY, Ahmed I, Iqbal M, Shafqat S, Anjum AA, et al. Description of novel variants in consanguineous Pakistani families affected with intellectual disability. Genes Genomics. 2022. Epub 20220212. doi: 10.1007/s13258-022-01219-y. PubMed PMID: 35150401.

291. Harripaul R, Vasli N, Mikhailov A, Rafiq MA, Mittal K, Windpassinger C, et al. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. Mol Psychiatry. 2018;23(4):973-84. Epub 20170411. doi: 10.1038/mp.2017.60. PubMed PMID: 28397838.

292. Mortreux J, Busa T, Germain DP, Nadeau G, Puechberty J, Coubes C, et al. The role of CNVs in the etiology of rare autosomal recessive disorders: the example of TRAPPC9-associated intellectual disability. Eur J Hum Genet. 2018;26(1):143-8. Epub 2017/11/29. doi: 10.1038/s41431-017-0018-x. PubMed PMID: 29187737; PubMed Central PMCID: 5838970.

293. Mir A, Kaufman L, Noor A, Motazacker MM, Jamil T, Azam M, et al. Identification of mutations in TRAPPC9, which encodes the NIK- and IKK-beta-binding protein, in nonsyndromic autosomal-recessive mental retardation. Am J Hum Genet. 2009;85(6):909-15. doi: 10.1016/j.ajhg.2009.11.009. PubMed PMID: 20004765; PubMed Central PMCID: 2790571.

294. Abbasi AA, Blaesius K, Hu H, Latif Z, Picker-Minh S, Khan MN, et al. Identification of a novel homozygous TRAPPC9 gene mutation causing non-syndromic intellectual disability, speech disorder, and secondary microcephaly. Am J Med Genet B Neuropsychiatr Genet. 2017;174(8):839-45. Epub 2017/10/14. doi: 10.1002/ajmg.b.32602. PubMed PMID: 29031008.

295. Mbimba T, Hussein NJ, Najeed A, Safadi FF. TRAPPC9: Novel insights into its trafficking and signaling pathways in health and disease (Review). Int J Mol Med. 2018;42(6):2991-7. Epub 2018/09/21. doi: 10.3892/ijmm.2018.3889. PubMed PMID: 30272317.

296. Mochida GH, Mahajnah M, Hill AD, Basel-Vanagaite L, Gleason D, Hill RS, et al. A truncating mutation of TRAPPC9 is associated with autosomal-recessive intellectual disability and postnatal microcephaly. Am J Hum Genet. 2009;85(6):897-902. doi: 10.1016/j.ajhg.2009.10.027. PubMed PMID: 20004763; PubMed Central PMCID: 2790576.

297. Kakar N, Goebel I, Daud S, Nürnberg G, Agha N, Ahmad A, et al. A homozygous splice site mutation in TRAPPC9 causes intellectual disability and microcephaly. Eur J Med Genet. 2012;55(12):727-31. Epub 20120830. doi: 10.1016/j.ejmg.2012.08.010. PubMed PMID: 22989526.

298. Marangi G, Leuzzi V, Manti F, Lattante S, Orteschi D, Pecile V, et al. TRAPPC9-related autosomal recessive intellectual disability: report of a new

mutation and clinical phenotype. Eur J Hum Genet. 2013;21(2):229-32. Epub 20120502. doi: 10.1038/ejhg.2012.79. PubMed PMID: 22549410; PubMed Central PMCID: 3548258.

299. Hnoonual A, Graidist P, Kritsaneepaiboon S, Limprasert P. Novel Compound Heterozygous Mutations in the TRAPPC9 Gene in Two Siblings With Autism and Intellectual Disability. Front Genet. 2019;10:61. Epub 2019/02/11. doi: 10.3389/fgene.2019.00061. PubMed PMID: 30853973; PubMed Central PMCID: 6396715.

300. Wilton KM, Gunderson LB, Hasadsri L, Wood CP, Schimmenti LA. Profound intellectual disability caused by homozygous TRAPPC9 pathogenic variant in a man from Malta. Mol Genet Genomic Med. 2020;8(5):e1211. Epub 20200311. doi: 10.1002/mgg3.1211. PubMed PMID: 32162493; PubMed Central PMCID: 7216808.

301. Alvarez-Mora MI, Corominas J, Gilissen C, Sanchez A, Madrigal I, Rodriguez-Revenga L. Novel Compound Heterozygous Mutation in TRAPPC9 Gene: The Relevance of Whole Genome Sequencing. Genes (Basel). 2021;12(4). Epub 20210412. doi: 10.3390/genes12040557. PubMed PMID: 33921338; PubMed Central PMCID: 8068822.

302. Krämer J, Beer M, Bode H, Winter B. Two Novel Compound Heterozygous Mutations in the TRAPPC9 Gene Reveal a Connection of Non-syndromic Intellectual Disability and Autism Spectrum Disorder. Front Genet. 2020;11:972. Epub 20210225. doi: 10.3389/fgene.2020.00972. PubMed PMID: 33719327; PubMed Central PMCID: 7947907.

303. Liang ZS, Cimino I, Yalcin B, Raghupathy N, Vancollie VE, Ibarra-Soria X, et al. Trappc9 deficiency causes parent-of-origin dependent microcephaly and obesity. PLoS Genet. 2020;16(9):e1008916. Epub 2020/09/02. doi: 10.1371/journal.pgen.1008916. PubMed PMID: 32877400; PubMed Central PMCID: 7467316.

304. Yousefipour F, Mozhdehipanah H, Mahjoubi F. Identification of two novel homozygous nonsense mutations in TRAPPC9 in two unrelated consanguineous families with intellectual Disability from Iran. Mol Genet Genomic Med. 2021;9(12):e1610. Epub 20210129. doi: 10.1002/mgg3.1610. PubMed PMID: 33513295; PubMed Central PMCID: 8683625.

305. Aslanger AD, Goncu B, Duzenli OF, Yucesan E, Sengenc E, Yesil G. Biallelic loss of TRAPPC9 function links vesicle trafficking pathway to autosomal recessive intellectual disability. J Hum Genet. 2022. Epub 20220105. doi: 10.1038/s10038-021-01007-8. PubMed PMID: 34983975.

306. Radenkovic S, Martinelli D, Zhang Y, Preston GJ, Maiorana A, Terracciano A, et al. TRAPPC9-CDG: A novel congenital disorder of glycosylation with dysmorphic features and intellectual disability. Genet Med. 2022. Epub 20220115. doi: 10.1016/j.gim.2021.12.012. PubMed PMID: 35042660.

307. Santos-Cortez RLP, Khan V, Khan FS, Mughal ZU, Chakchouk I, Lee K, et al. Novel candidate genes and variants underlying autosomal recessive

neurodevelopmental disorders with intellectual disability. Hum Genet. 2018;137(9):735-52. Epub 2018/08/22. doi: 10.1007/s00439-018-1928-6. PubMed PMID: 30167849; PubMed Central PMCID: 6201268.

308. Bögershausen N, Shahrzad N, Chong JX, von Kleist-Retzow JC, Stanga D, Li Y, et al. Recessive TRAPPC11 mutations cause a disease spectrum of limb girdle muscular dystrophy and myopathy with movement disorder and intellectual disability. Am J Hum Genet. 2013;93(1):181-90. Epub 2013/07/03. doi: 10.1016/j.ajhg.2013.05.028. PubMed PMID: 23830518; PubMed Central PMCID: 3710757.

309. Liang WC, Zhu W, Mitsuhashi S, Noguchi S, Sacher M, Ogawa M, et al. Congenital muscular dystrophy with fatty liver and infantile-onset cataract caused by TRAPPC11 mutations: broadening of the phenotype. Skelet Muscle. 2015;5:29. Epub 2015/08/28. doi: 10.1186/s13395-015-0056-4. PubMed PMID: 26322222; PubMed Central PMCID: 4551700.

310. Koehler K, Milev MP, Prematilake K, Reschke F, Kutzner S, Jühlen R, et al. A novel *TRAPPC11* mutation in two Turkish families associated with cerebral atrophy, global retardation, scoliosis, achalasia and alacrima. J Med Genet. 2017;54(3):176-85. Epub 2016/10/05. doi: 10.1136/jmedgenet-2016-104108. PubMed PMID: 27707803.

311. Milev MP, Stanga D, Schänzer A, Nascimento A, Saint-Dic D, Ortez C, et al. Characterization of three TRAPPC11 variants suggests a critical role for the extreme carboxy terminus of the protein. Sci Rep. 2019;9(1):14036. Epub 2019/10/01. doi: 10.1038/s41598-019-50415-6. PubMed PMID: 31575891; PubMed Central PMCID: 6773699.

312. Munot P, McCrea N, Torelli S, Manzur A, Sewry C, Chambers D, et al. TRAPPC11-related muscular dystrophy with hypoglycosylation of alpha-dystroglycan in skeletal muscle and brain. Neuropathol Appl Neurobiol. 2022;48(2):e12771. Epub 20211111. doi: 10.1111/nan.12771. PubMed PMID: 34648194.

313. Hadouiri N, Thomas Q, Darmency V, Dulieu V, De Rougemont MM, Bruel AL, et al. Homozygous TRAPPC11 truncating variant revealing segmental uniparental disomy of chromosome 4 as a cause of a recessive limb-girdle muscular dystrophy-18. Clin Genet. 2021;100(5):643-4. Epub 20210825. doi: 10.1111/cge.14045. PubMed PMID: 34435357.

314. Wang X, Wu Y, Cui Y, Wang N, Folkersen L, Wang Y. Novel TRAPPC11 Mutations in a Chinese Pedigree of Limb Girdle Muscular Dystrophy. Case Rep Genet. 2018;2018:8090797. Epub 20180716. doi: 10.1155/2018/8090797. PubMed PMID: 30105108; PubMed Central PMCID: 6076900.

315. Matalonga L, Bravo M, Serra-Peinado C, García-Pelegrí E, Ugarteburu O, Vidal S, et al. Mutations in TRAPPC11 are associated with a congenital disorder of glycosylation. Hum Mutat. 2017;38(2):148-51. Epub 20161126. doi: 10.1002/humu.23145. PubMed PMID: 27862579.

316. Fee DB, Harmelink M, Monrad P, Pyzik E. Siblings With Mutations in TRAPPC11 Presenting With Limb-Girdle Muscular Dystrophy 2S. J Clin Neuromuscul Dis. 2017;19(1):27-30. doi: 10.1097/CND.0000000000000173. PubMed PMID: 28827486.

317. Milev MP, Grout ME, Saint-Dic D, Cheng YH, Glass IA, Hale CJ, et al. Mutations in TRAPPC12 Manifest in Progressive Childhood Encephalopathy and Golgi Dysfunction. Am J Hum Genet. 2017;101(2):291-9. doi: 10.1016/j.ajhg.2017.07.006. PubMed PMID: 28777934; PubMed Central PMCID: 5544387.

318. Aslanger AD, Demiral E, Sonmez-Sahin S, Guler S, Goncu B, Yucesan E, et al. Expanding Clinical Phenotype of TRAPPC12-Related Childhood Encephalopathy: Two Cases and Review of Literature. Neuropediatrics. 2020;51(6):430-4. Epub 20200505. doi: 10.1055/s-0040-1710526. PubMed PMID: 32369837.

319. Gass JM, Head BB, Shields SM, Stevenson RE, Louie RJ. Hydrocephaly associated with compound heterozygous alterations in TRAPPC12. Birth Defects Res. 2020;112(13):1028-34. Epub 20200429. doi: 10.1002/bdr2.1699. PubMed PMID: 32347653.

320. Perez Y, Bar-Yaacov R, Kadir R, Wormser O, Shelef I, Birk OS, et al. Mutations in the microtubule-associated protein MAP11 (C7orf43) cause microcephaly in humans and zebrafish. Brain. 2019;142(3):574-85. doi: 10.1093/brain/awz004. PubMed PMID: 30715179; PubMed Central PMCID: 6391606.

321. Yarwood R, Hellicar J, Woodman PG, Lowe M. Membrane trafficking in health and disease. Dis Model Mech. 2020;13(4). Epub 20200430. doi: 10.1242/dmm.043448. PubMed PMID: 32433026; PubMed Central PMCID: 7197876.

322. Khan V. Identification of genes responsible for intellectual disability in consanguineous families: Quaid-i-Azam University, Islamabad; 2019.

323. Scales SJ, Pepperkok R, Kreis TE. Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. Cell. 1997;90(6):1137-48. PubMed PMID: 9323141.

324. Mikhaleva A, Kannan M, Wagner C, Yalcin B. Histomorphological Phenotyping of the Adult Mouse Brain. Curr Protoc Mouse Biol. 2016;6(3):307-32. Epub 2016/09/01. doi: 10.1002/cpmo.12. PubMed PMID: 27584555.

325. Rawlins LE, Almousa H, Khan S, Collins SC, Milev MP, Leslie J, et al. Biallelic variants in TRAPPC10 cause a microcephalic TRAPPopathy disorder in humans and mice. PLoS Genet. 2022;18(3):e1010114. Epub 20220317. doi: 10.1371/journal.pgen.1010114. PubMed PMID: 35298461; PubMed Central PMCID: 8963566.

326. Vazquez-Martinez R, Cruz-Garcia D, Duran-Prado M, Peinado JR, Castaño JP, Malagon MM. Rab18 inhibits secretory activity in neuroendocrine

cells by interacting with secretory granules. Traffic. 2007;8(7):867-82. Epub 2007/05/04. doi: 10.1111/j.1600-0854.2007.00570.x. PubMed PMID: 17488286.

327.   Bem D, Yoshimura S, Nunes-Bastos R, Bond FC, Bond FF, Kurian MA, et al. Loss-of-function mutations in RAB18 cause Warburg micro syndrome. Am J Hum Genet. 2011;88(4):499-507. doi: 10.1016/j.ajhg.2011.03.012. PubMed PMID: 21473985; PubMed Central PMCID: 3071920.

328.   Winkle CC, Gupton SL. Membrane Trafficking in Neuronal Development: Ins and Outs of Neural Connectivity. Int Rev Cell Mol Biol. 2016;322:247-80. Epub 20160106. doi: 10.1016/bs.ircmb.2015.10.003. PubMed PMID: 26940520; PubMed Central PMCID: 4797321.

329.   Ethell IM, Hagihara K, Miura Y, Irie F, Yamaguchi Y. Synbindin, A novel syndecan-2-binding protein in neuronal dendritic spines. J Cell Biol. 2000;151(1):53-68. doi: 10.1083/jcb.151.1.53. PubMed PMID: 11018053; PubMed Central PMCID: 2189810.

330.   Bodnar B, DeGruttola A, Zhu Y, Lin Y, Zhang Y, Mo X, et al. Emerging role of NIK/IKK2-binding protein (NIBP)/trafficking protein particle complex 9 (TRAPPC9) in nervous system diseases. Transl Res. 2020;224:55-70. Epub 2020/05/17. doi: 10.1016/j.trsl.2020.05.001. PubMed PMID: 32434006; PubMed Central PMCID: 7442628.

331.   Ke Y, Weng M, Chhetri G, Usman M, Li Y, Yu Q, et al. Trappc9 deficiency in mice impairs learning and memory by causing imbalance of dopamine D1 and D2 neurons. Sci Adv. 2020;6(47). Epub 2020/11/18. doi: 10.1126/sciadv.abb7781. PubMed PMID: 33208359; PubMed Central PMCID: 7673810.

332.   Zhang Y, Liu S, Wang H, Yang W, Li F, Yang F, et al. Elevated NIBP/TRAPPC9 mediates tumorigenesis of cancer cells through NFκB signaling. Oncotarget. 2015;6(8):6160-78. doi: 10.18632/oncotarget.3349. PubMed PMID: 25704885; PubMed Central PMCID: 4467429.

333.   Baple EL, Maroofian R, Chioza BA, Izadi M, Cross HE, Al-Turki S, et al. Mutations in KPTN cause macrocephaly, neurodevelopmental delay, and seizures. Am J Hum Genet. 2014;94(1):87-94. Epub 2013/11/14. doi: 10.1016/j.ajhg.2013.10.001. PubMed PMID: 24239382; PubMed Central PMCID: 3882725.

334.   Pajusalu S, Reimand T, Ounap K. Novel homozygous mutation in KPTN gene causing a familial intellectual disability-macrocephaly syndrome. American journal of medical genetics Part A. 2015;167a(8):1913-5. Epub 2015/04/08. doi: 10.1002/ajmg.a.37105

10.1002/ajmg.a.37105. Epub 2015 Apr 5. PubMed PMID: 25847626.

335.   Thiffault I, Atherton A, Heese BA, Abdelmoity A, Pawar K, Farrow E, et al. Pathogenic variants in KPTN gene identified by Clinical Whole Whole-Genome Sequencing. Journal of Genetic and Hereditary Case Studies. 2018;1(1):1-6.

336.   Lucena PH, Armani-Franceschi G, Bispo-Torres AC, Bandeira ID, Lucena MFG, Maldonado I, et al. KPTN gene homozygous variant-related syndrome in the northeast of Brazil: A case report. Am J Med Genet A. 2020;182(4):762-7. Epub 2020/01/30. doi: 10.1002/ajmg.a.61492. PubMed PMID: 31999056.

337.   Pacio Miguez M, Santos-Simarro F, García-Miñaúr S, Velázquez Fragua R, Del Pozo Á, Solís M, et al. Pathogenic variants in KPTN, a rare cause of macrocephaly and intellectual disability. Am J Med Genet A. 2020;182(10):2222-5. Epub 2020/08/18. doi: 10.1002/ajmg.a.61778. PubMed PMID: 32808430.

338.   Thiffault I, Atherton A, Heese BA, Abdelmoity A, Pawar K, Farrow E, et al. Pathogenic variants in KPTN gene identified by Clinical Whole Whole-Genome Sequencing. Journal of Genetic and Hereditary Case Studies. 2018;1(1):1-6.

339.   Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med. 2016;18(7):696-704. Epub 2015/12/04. doi: 10.1038/gim.2015.148

10.1038/gim.2015.148. Epub 2015 Dec 3. PubMed PMID: 26633542.

340.   Bearer EL, Chen AF, Chen AH, Li Z, Mark HF, Smith RJ, et al. 2E4/Kaptin (KPTN)--a candidate gene for the hearing loss locus, DFNA4. Ann Hum Genet. 2000;64(Pt 3):189-96. PubMed PMID: 11409409; PubMed Central PMCID: 3376086.

341.   Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. Nucleic Acids Res. 2021;49(D1):D344-D54. doi: 10.1093/nar/gkaa977. PubMed PMID: 33156333.

342.   Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. Nucleic Acids Res. 2018;46(D1):D754-D61. doi: 10.1093/nar/gkx1098. PubMed PMID: 29155950; PubMed Central PMCID: 5753206.

343.   Bearer EL. An actin-associated protein present in the microtubule organizing center and the growth cones of PC-12 cells. J Neurosci. 1992;12(3):750-61. Epub 1992/03/01. PubMed PMID: 1372044.

344.   Bearer EL, Abraham MT. 2E4 (kaptin): a novel actin-associated protein from human blood platelets found in lamellipodia and the tips of the stereocilia of the inner ear. European journal of cell biology. 1999;78(2):117-26. Epub 1999/04/01. doi: 10.1016/s0171-9335(99)80013-2

10.1016/S0171-9335(99)80013-2. PubMed PMID: 10099934.

345.   Dent EW, Merriam EB, Hu X. The dynamic cytoskeleton: backbone of dendritic spine plasticity. Curr Opin Neurobiol. 2011;21(1):175-81. Epub 2010/09/09. doi: 10.1016/j.conb.2010.08.013. PubMed PMID: 20832290; PubMed Central PMCID: 3010448.

346.   Kessels MM, Schwintzer L, Schlobinski D, Qualmann B. Controlling actin cytoskeletal organization and dynamics during neuronal morphogenesis.

European journal of cell biology. 2011;90(11):926-33. Epub 2010/10/20. doi: 10.1016/j.ejcb.2010.08.011. PubMed PMID: 20965607.

347.   Svitkina T, Lin WH, Webb DJ, Yasuda R, Wayman GA, Van Aelst L, et al. Regulation of the postsynaptic cytoskeleton: roles in development, plasticity, and disorders. J Neurosci. 2010;30(45):14937-42. doi: 10.1523/JNEUROSCI.4276-10.2010. PubMed PMID: 21068295; PubMed Central PMCID: 3074493.

348.   Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'Elia E, et al. Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. Am J Hum Genet. 2010;86(2):185-95. doi: 10.1016/j.ajhg.2010.01.011. PubMed PMID: 20159109; PubMed Central PMCID: 2820185.

349.   von Bohlen Und Halbach O. Dendritic spine abnormalities in mental retardation. Cell Tissue Res. 2010;342(3):317-23. Epub 2010/11/16. doi: 10.1007/s00441-010-1070-9. PubMed PMID: 21080001.

350.   Piard J, Hu JH, Campeau PM, Rzonca S, Van Esch H, Vincent E, et al. FRMPD4 mutations cause X-linked intellectual disability and disrupt dendritic spine morphogenesis. Hum Mol Genet. 2018;27(4):589-600. doi: 10.1093/hmg/ddx426. PubMed PMID: 29267967; PubMed Central PMCID: 5886117.

351.   Lipton JO, Sahin M. The neurology of mTOR. Neuron. 2014;84(2):275-91. Epub 2014/10/22. doi: 10.1016/j.neuron.2014.09.034. PubMed PMID: 25374355; PubMed Central PMCID: 4223653.

352.   Saxton RA, Sabatini DM. mTOR Signaling in Growth, Metabolism, and Disease. Cell. 2017;169(2):361-71. doi: 10.1016/j.cell.2017.03.035. PubMed PMID: 28388417.

353.   Hoeffer CA, Klann E. mTOR signaling: at the crossroads of plasticity, memory and disease. Trends Neurosci. 2010;33(2):67-75. Epub 20091204. doi: 10.1016/j.tins.2009.11.003. PubMed PMID: 19963289; PubMed Central PMCID: 2821969.

354.   Johnson SC, Rabinovitch PS, Kaeberlein M. mTOR is a key modulator of ageing and age-related disease. Nature. 2013;493(7432):338-45. doi: 10.1038/nature11861. PubMed PMID: 23325216; PubMed Central PMCID: 3687363.

355.   Frankel WN, Yang Y, Mahaffey CL, Beyer BJ, O'Brien TP. Szt2, a novel gene for seizure threshold in mice. Genes Brain Behav. 2009;8(5):568-76. doi: 10.1111/j.1601-183X.2009.00509.x. PubMed PMID: 19624305; PubMed Central PMCID: 2728062.

356.   Basel-Vanagaite L, Hershkovitz T, Heyman E, Raspall-Chaure M, Kakar N, Smirin-Yosef P, et al. Biallelic SZT2 Mutations Cause Infantile Encephalopathy with Epilepsy and Dysmorphic Corpus Callosum. Am J Hum Genet. 932013. p. 524-9.

357.   Negishi Y, Miya F, Hattori A, Johmura Y, Nakagawa M, Ando N, et al. A combination of genetic and biochemical analyses for the diagnosis of PI3K-AKT-mTOR pathway-associated megalencephaly. BMC Med Genet. 2017;18(1):4. Epub 20170113. doi: 10.1186/s12881-016-0363-6. PubMed PMID: 28086757; PubMed Central PMCID: 5237172.

358.   Nakamura Y, Togawa Y, Okuno Y, Muramatsu H, Nakabayashi K, Kuroki Y, et al. Biallelic mutations in SZT2 cause a discernible clinical entity with epilepsy, developmental delay, macrocephaly and a dysmorphic corpus callosum. Brain Dev. 2018;40(2):134-9. Epub 2017/10/12. doi: 10.1016/j.braindev.2017.08.003. PubMed PMID: 28893434.

359.   Dang LT, Glanowska KM, Iffland Ii PH, Barnes AE, Baybis M, Liu Y, et al. Multimodal Analysis of STRADA Function in Brain Development. Front Cell Neurosci. 2020;14:122. Epub 20200508. doi: 10.3389/fncel.2020.00122. PubMed PMID: 32457579; PubMed Central PMCID: 7227375.

360.   Hartman NW, Lin TV, Zhang L, Paquelet GE, Feliciano DM, Bordey A. mTORC1 targets the translational repressor 4E-BP2, but not S6 kinase 1/2, to regulate neural stem cell self-renewal in vivo. Cell Rep. 2013;5(2):433-44. Epub 20131017. doi: 10.1016/j.celrep.2013.09.017. PubMed PMID: 24139800.

361.   Ka M, Condorelli G, Woodgett JR, Kim WY. mTOR regulates brain morphogenesis by mediating GSK3 signaling. Development. 2014;141(21):4076-86. Epub 20141001. doi: 10.1242/dev.108282. PubMed PMID: 25273085; PubMed Central PMCID: 4302893.

362.   Meng D, Frank AR, Jewell JL. mTOR signaling in stem and progenitor cells. Development. 2018;145(1). Epub 20180108. doi: 10.1242/dev.152595. PubMed PMID: 29311260; PubMed Central PMCID: 5825873.

363.   Pollen AA, Bhaduri A, Andrews MG, Nowakowski TJ, Meyerson OS, Mostajo-Radji MA, et al. Establishing Cerebral Organoids as Models of Human-Specific Brain Evolution. Cell. 2019;176(4):743-56.e17. doi: 10.1016/j.cell.2019.01.017. PubMed PMID: 30735633; PubMed Central PMCID: 6544371.

364.   Lee JH, Huynh M, Silhavy JL, Kim S, Dixon-Salazar T, Heiberg A, et al. De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. Nat Genet. 2012;44(8):941-5. Epub 2012/06/24. doi: 10.1038/ng.2329. PubMed PMID: 22729223; PubMed Central PMCID: 4417942.

365.   Gordo G, Tenorio J, Arias P, Santos-Simarro F, García-Miñaur S, Moreno JC, et al. mTOR mutations in Smith-Kingsmore syndrome: four additional patients and a review. Clin Genet. 2017. Epub 2017/09/11. doi: 10.1111/cge.13135. PubMed PMID: 28892148.

366.   D'Gama AM, Geng Y, Couto JA, Martin B, Boyle EA, LaCoursiere CM, et al. Mammalian target of rapamycin pathway mutations cause hemimegalencephaly and focal cortical dysplasia. Ann Neurol. 2015;77(4):720-

5. Epub 20150226. doi: 10.1002/ana.24357. PubMed PMID: 25599672; PubMed Central PMCID: 4471336.

367.   Wolfson RL, Chantranupong L, Wyant GA, Gu X, Orozco JM, Shen K, et al. KICSTOR recruits GATOR1 to the lysosome and is necessary for nutrients to regulate mTORC1. Nature. 2017;543(7645):438-42. Epub 2017/02/15. doi: 10.1038/nature21423. PubMed PMID: 28199306; PubMed Central PMCID: 5360989.

368.   Bar-Peled L, Chantranupong L, Cherniack AD, Chen WW, Ottina KA, Grabiner BC, et al. A Tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. Science. 2013;340(6136):1100-6. doi: 10.1126/science.1232044. PubMed PMID: 23723238; PubMed Central PMCID: 3728654.

369.   Parker WE, Orlova KA, Parker WH, Birnbaum JF, Krymskaya VP, Goncharov DA, et al. Rapamycin prevents seizures after depletion of STRADA in a rare neurodevelopmental disorder. Sci Transl Med. 2013;5(182):182ra53. doi: 10.1126/scitranslmed.3005271. PubMed PMID: 23616120; PubMed Central PMCID: 3720125.

370.   Nelson K, Jackman C, Bell J, Shih CS, Payne K, Dlouhy S, et al. Novel Homozygous Deletion in STRADA Gene Associated With Polyhydramnios, Megalencephaly, and Epilepsy in 2 Siblings: Implications for Diagnosis and Treatment. J Child Neurol. 2018;33(14):925-9. Epub 2018/10/12. doi: 10.1177/0883073818802724. PubMed PMID: 30311510.

371.   Li M, Zhou Y, Chen C, Yang T, Zhou S, Chen S, et al. Efficacy and safety of mTOR inhibitors (rapamycin and its analogues) for tuberous sclerosis complex: a meta-analysis. Orphanet J Rare Dis. 2019;14(1):39. Epub 2019/02/13. doi: 10.1186/s13023-019-1012-x. PubMed PMID: 30760308; PubMed Central PMCID: 6373010.

372.   Zak M, Ledbetter M, Maertens P. Infantile Lhermitte-Duclos Disease Treated Successfully With Rapamycin. J Child Neurol. 2017;32(3):322-6. Epub 2016/12/08. doi: 10.1177/0883073816681340. PubMed PMID: 27932596.

373.   Schmid GL, Kässner F, Uhlig HH, Körner A, Kratzsch J, Händel N, et al. Sirolimus treatment of severe PTEN hamartoma tumor syndrome: case report and in vitro studies. Pediatr Res. 2014;75(4):527-34. Epub 2013/12/23. doi: 10.1038/pr.2013.246. PubMed PMID: 24366516.

374.   Iacobas I, Burrows PE, Adams DM, Sutton VR, Hollier LH, Chintagumpala MM. Oral rapamycin in the treatment of patients with hamartoma syndromes and PTEN mutation. Pediatr Blood Cancer. 2011;57(2):321-3. Epub 2011/02/25. doi: 10.1002/pbc.23098. PubMed PMID: 21360661.

375.   Baldassari S, Licchetta L, Tinuper P, Bisulli F, Pippucci T. GATOR1 complex: the common genetic actor in focal epilepsies. J Med Genet. 2016;53(8):503-10. Epub 2016/05/19. doi: 10.1136/jmedgenet-2016-103883. PubMed PMID: 27208208.

376. Pizzino A, Whitehead M, Sabet Rasekh P, Murphy J, Helman G, Bloom M, et al. Mutations in SZT2 result in early-onset epileptic encephalopathy and leukoencephalopathy. Am J Med Genet A. 2018;176(6):1443-8. Epub 2018/04/25. doi: 10.1002/ajmg.a.38717. PubMed PMID: 29696782.

377. Bi W, Glass IA, Muzny DM, Gibbs RA, Eng CM, Yang Y, et al. Whole exome sequencing identifies the first STRADA point mutation in a patient with polyhydramnios, megalencephaly, and symptomatic epilepsy syndrome (PMSE). Am J Med Genet A. 2016;170(8):2181-5. Epub 2016/05/12. doi: 10.1002/ajmg.a.37727. PubMed PMID: 27170158.

378. Baldassari S, Picard F, Verbeek NE, van Kempen M, Brilstra EH, Lesca G, et al. The landscape of epilepsy-related GATOR1 variants. Genet Med. 2018. Epub 2018/08/10. doi: 10.1038/s41436-018-0060-2. PubMed PMID: 30093711; PubMed Central PMCID: 6292495.

379. Iffland PH, Carson V, Bordey A, Crino PB. GATORopathies: The role of amino acid regulatory gene mutations in epilepsy and cortical malformations. Epilepsia. 2019;60(11):2163-73. Epub 20191017. doi: 10.1111/epi.16370. PubMed PMID: 31625153; PubMed Central PMCID: 7155771.

380. White JK, Gerdin AK, Karp NA, Ryder E, Buljan M, Bussell JN, et al. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell. 2013;154(2):452-64. doi: 10.1016/j.cell.2013.06.022. PubMed PMID: 23870131; PubMed Central PMCID: 3717207.

381. Seibenhener ML, Wooten MC. Use of the Open Field Maze to measure locomotor and anxiety-like behavior in mice. J Vis Exp. 2015;(96):e52434. Epub 20150206. doi: 10.3791/52434. PubMed PMID: 25742564; PubMed Central PMCID: 4354627.

382. Levitin M. Applying animal modelling to understand rare novel neurodevelopmental disorders associated with intellectual disability: University of Cambridge; 2017.

383. Crawley J, Goodwin FK. Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines. Pharmacol Biochem Behav. 1980;13(2):167-70. doi: 10.1016/0091-3057(80)90067-2. PubMed PMID: 6106204.

384. Bourin M, Hascoët M. The mouse light/dark box test. Eur J Pharmacol. 2003;463(1-3):55-65. doi: 10.1016/s0014-2999(03)01274-3. PubMed PMID: 12600702.

385. Sánchez-Andrade G, James BM, Kendrick KM. Neural encoding of olfactory recognition memory. J Reprod Dev. 2005;51(5):547-58. doi: 10.1262/jrd.17031. PubMed PMID: 16284449.

386. Ferguson JN, Aldag JM, Insel TR, Young LJ. Oxytocin in the medial amygdala is essential for social recognition in the mouse. J Neurosci.

2001;21(20):8278-85. PubMed PMID: 11588199; PubMed Central PMCID: 6763861.

387.   Winslow JT, Insel TR. Neuroendocrine basis of social recognition. Curr Opin Neurobiol. 2004;14(2):248-53. doi: 10.1016/j.conb.2004.03.009. PubMed PMID: 15082332.

388.   Koopmans G, Blokland A, van Nieuwenhuijzen P, Prickaerts J. Assessment of spatial learning abilities of mice in a new circular maze. Physiol Behav. 2003;79(4-5):683-93. doi: 10.1016/s0031-9384(03)00171-9. PubMed PMID: 12954410.

389.   Bussey TJ, Holmes A, Lyon L, Mar AC, McAllister KA, Nithianantharajah J, et al. New translational assays for preclinical modelling of cognition in schizophrenia: the touchscreen testing method for mice and rats. Neuropharmacology. 2012;62(3):1191-203. Epub 20110421. doi: 10.1016/j.neuropharm.2011.04.011. PubMed PMID: 21530550; PubMed Central PMCID: 3168710.

390.   de Carlos F, Alvarez-Suárez A, Costilla S, Noval I, Vega J, Cobo J. 3D-µCT Cephalometric Measurements in Mice. In: Saba L, editor. Computed Tomography: Special applications: InTechOpen; 2011.

391.   Xu H, Lenhart SA, Chu EY, Chavez MB, Wimer HF, Dimori M, et al. Dental and craniofacial defects in the Crtap$^{-/-}$ mouse model of osteogenesis imperfecta

type VII. Dev Dyn. 2020;249(7):884-97. Epub 20200312. doi: 10.1002/dvdy.166. PubMed PMID: 32133710; PubMed Central PMCID: 7727892.

392.   Baddam P, Biancardi V, Roth DM, Eaton F, Thereza-Bussolaro C, Mandal R, et al. Neural crest-specific loss of Bmp7 leads to midfacial hypoplasia, nasal airway obstruction and disordered breathing, modeling obstructive sleep apnea. Dis Model Mech. 2021. Epub 20210111. doi: 10.1242/dmm.047738. PubMed PMID: 33431521; PubMed Central PMCID: 7888714.

393.   Cole TJ. The LMS method for constructing normalized growth standards. Eur J Clin Nutr. 1990;44(1):45-60. PubMed PMID: 2354692.

394.   Yarkoni T, Speer NK, Zacks JM. Neural substrates of narrative comprehension and memory. Neuroimage. 2008;41(4):1408-25. Epub 20080415. doi: 10.1016/j.neuroimage.2008.03.062. PubMed PMID: 18499478; PubMed Central PMCID: 2580728.

395.   Ruvinsky I, Meyuhas O. Ribosomal protein S6 phosphorylation: from protein synthesis to cell size. Trends Biochem Sci. 2006;31(6):342-8. Epub 20060506. doi: 10.1016/j.tibs.2006.04.003. PubMed PMID: 16679021.

396.   Biever A, Valjent E, Puighermanal E. Ribosomal Protein S6 Phosphorylation in the Nervous System: From Regulation to Function. Front Mol Neurosci. 2015;8:75. Epub 20151216. doi: 10.3389/fnmol.2015.00075. PubMed PMID: 26733799; PubMed Central PMCID: 4679984.

397.	Ballou LM, Lin RZ. Rapamycin and mTOR kinase inhibitors. J Chem Biol. 2008;1(1-4):27-36. Epub 20080515. doi: 10.1007/s12154-008-0003-5. PubMed PMID: 19568796; PubMed Central PMCID: 2698317.

398.	Spilman P, Podlutskaya N, Hart MJ, Debnath J, Gorostiza O, Bredesen D, et al. Inhibition of mTOR by rapamycin abolishes cognitive deficits and reduces amyloid-beta levels in a mouse model of Alzheimer's disease. PLoS One. 2010;5(4):e9979. Epub 20100401. doi: 10.1371/journal.pone.0009979. PubMed PMID: 20376313; PubMed Central PMCID: 2848616.

399.	Way SW, Rozas NS, Wu HC, McKenna J, Reith RM, Hashmi SS, et al. The differential effects of prenatal and/or postnatal rapamycin on neurodevelopmental defects and cognition in a neuroglial mouse model of tuberous sclerosis complex. Hum Mol Genet. 2012;21(14):3226-36. Epub 20120424. doi: 10.1093/hmg/dds156. PubMed PMID: 22532572; PubMed Central PMCID: 3384384.

400.	Qi Y, Zhang XJ, Renier N, Wu Z, Atkin T, Sun Z, et al. Combined small-molecule inhibition accelerates the derivation of functional cortical neurons from human pluripotent stem cells. Nat Biotechnol. 2017;35(2):154-63. Epub 20170123. doi: 10.1038/nbt.3777. PubMed PMID: 28112759; PubMed Central PMCID: 5516899.

401.	Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24. Epub 2015/03/05. doi: 10.1038/gim.2015.30. PubMed PMID: 25741868; PubMed Central PMCID: 4544753.

402.	Ellard S, Baple E, Callaway A, Berry I, Forrester N, Turnbull C, et al. ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. Association for Clinical Genomic Science; 2020.

403.	Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex. 2009;19(12):2767-96. Epub 2009/03/27. doi: 10.1093/cercor/bhp055. PubMed PMID: 19329570; PubMed Central PMCID: 2774390.

404.	Lim JS, Lee JH. Brain somatic mutations in MTOR leading to focal cortical dysplasia. BMB Rep. 2016;49(2):71-2. PubMed PMID: 26779999; PubMed Central PMCID: 4915119.

405.	Poduri A, Evrony GD, Cai X, Elhosary PC, Beroukhim R, Lehtinen MK, et al. Somatic activation of AKT3 causes hemispheric developmental brain malformations. Neuron. 2012;74(1):41-8. doi: 10.1016/j.neuron.2012.03.010. PubMed PMID: 22500628; PubMed Central PMCID: 3460551.

406.	Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456

humans. Nature. 2020;581(7809):434-43. Epub 20200527. doi: 10.1038/s41586-020-2308-7. PubMed PMID: 32461654; PubMed Central PMCID: 7334197.

407. Reijnders MRF, Kousi M, van Woerden GM, Klein M, Bralten J, Mancini GMS, et al. Variation in a range of mTOR-related genes associates with intracranial volume and intellectual disability. Nat Commun. 2017;8(1):1052. Epub 20171020. doi: 10.1038/s41467-017-00933-6. PubMed PMID: 29051493; PubMed Central PMCID: 5648772.

408. Crino PB. The mTOR signalling cascade: paving new roads to cure neurological disease. Nat Rev Neurol. 2016;12(7):379-92. Epub 2016/06/24. doi: 10.1038/nrneurol.2016.81. PubMed PMID: 27340022.

409. Tatton-Brown K, Weksberg R. Molecular mechanisms of childhood overgrowth. Am J Med Genet C Semin Med Genet. 2013;163C(2):71-5. Epub 20130418. doi: 10.1002/ajmg.c.31362. PubMed PMID: 23606607.

410. Keppler-Noreuil KM, Parker VE, Darling TN, Martinez-Agosto JA. Somatic overgrowth disorders of the PI3K/AKT/mTOR pathway & therapeutic strategies. Am J Med Genet C Semin Med Genet. 2016;172(4):402-21. Epub 20161118. doi: 10.1002/ajmg.c.31531. PubMed PMID: 27860216; PubMed Central PMCID: 5592089.

411. Peng M, Yin N, Li MO. SZT2 dictates GATOR control of mTORC1 signalling. Nature. 2017;543(7645):433-7. Epub 2017/02/15. doi: 10.1038/nature21378. PubMed PMID: 28199315; PubMed Central PMCID: 5570594.

412. Overwater IE, Rietman AB, Bindels-de Heus K, Looman CW, Rizopoulos D, Sibindi TM, et al. Sirolimus for epilepsy in children with tuberous sclerosis complex: A randomized controlled trial. Neurology. 2016;87(10):1011-8. Epub 2016/08/10. doi: 10.1212/WNL.0000000000003077. PubMed PMID: 27511181.

413. French JA, Lawson JA, Yapici Z, Ikeda H, Polster T, Nabbout R, et al. Adjunctive everolimus therapy for treatment-resistant focal-onset seizures associated with tuberous sclerosis (EXIST-3): a phase 3, randomised, double-blind, placebo-controlled study. Lancet. 2016;388(10056):2153-63. Epub 2016/09/06. doi: 10.1016/S0140-6736(16)31419-2. PubMed PMID: 27613521.

414. Curatolo P. Mechanistic target of rapamycin (mTOR) in tuberous sclerosis complex-associated epilepsy. Pediatr Neurol. 2015;52(3):281-9. Epub 2014/11/20. doi: 10.1016/j.pediatrneurol.2014.10.028. PubMed PMID: 25591831.

415. Bissler JJ, Kingswood JC, Radzikowska E, Zonnenberg BA, Belousova E, Frost MD, et al. Everolimus long-term use in patients with tuberous sclerosis complex: Four-year update of the EXIST-2 study. PLoS One. 2017;12(8):e0180939. Epub 2017/08/09. doi: 10.1371/journal.pone.0180939. PubMed PMID: 28792952; PubMed Central PMCID: 5549893.

416.   Chen Y, Zhou X. Research progress of mTOR inhibitors. Eur J Med Chem. 2020;208:112820. Epub 20200913. doi: 10.1016/j.ejmech.2020.112820. PubMed PMID: 32966896.

417.   Xu T, Sun D, Chen Y, Ouyang L. Targeting mTOR for fighting diseases: A revisited review of mTOR inhibitors. Eur J Med Chem. 2020;199:112391. Epub 20200504. doi: 10.1016/j.ejmech.2020.112391. PubMed PMID: 32416459.

418.   Kwon CH, Zhu X, Zhang J, Knoop LL, Tharp R, Smeyne RJ, et al. Pten regulates neuronal soma size: a mouse model of Lhermitte-Duclos disease. Nat Genet. 2001;29(4):404-11. doi: 10.1038/ng781. PubMed PMID: 11726927.

419.   Ehninger D, Han S, Shilyansky C, Zhou Y, Li W, Kwiatkowski DJ, et al. Reversal of learning deficits in a Tsc2+/- mouse model of tuberous sclerosis. Nat Med. 2008;14(8):843-8. Epub 20080622. doi: 10.1038/nm1788. PubMed PMID: 18568033; PubMed Central PMCID: 2664098.

420.   Ehninger D, Li W, Fox K, Stryker MP, Silva AJ. Reversing neurodevelopmental disorders in adults. Neuron. 2008;60(6):950-60. doi: 10.1016/j.neuron.2008.12.007. PubMed PMID: 19109903; PubMed Central PMCID: 2710296.

421.   Richter T, Nestler-Parr S, Babela R, Khan ZM, Tesoro T, Molsen E, et al. Rare Disease Terminology and Definitions-A Systematic Global Review: Report of the ISPOR Rare Disease Special Interest Group. Value Health. 2015;18(6):906-14. Epub 20150818. doi: 10.1016/j.jval.2015.05.008. PubMed PMID: 26409619.

422.   Savatt JM, Myers SM. Genetic Testing in Neurodevelopmental Disorders. Front Pediatr. 2021;9:526779. Epub 20210219. doi: 10.3389/fped.2021.526779. PubMed PMID: 33681094; PubMed Central PMCID: 7933797.

423.   Tărlungeanu DC, Novarino G. Genomics in neurodevelopmental disorders: an avenue to personalized medicine. Exp Mol Med. 2018;50(8):1-7. Epub 20180807. doi: 10.1038/s12276-018-0129-7. PubMed PMID: 30089840; PubMed Central PMCID: 6082867.

424.   Fasham J, Leslie JS, Harrison JW, Deline J, Williams KB, Kuhl A, et al. No association between SCN9A and monogenic human epilepsy disorders. PLoS Genet. 2020;16(11):e1009161. Epub 20201120. doi: 10.1371/journal.pgen.1009161. PubMed PMID: 33216760; PubMed Central PMCID: 7717534.

425.   Patton J, Brewer C, Chien W, Johnston JJ, Griffith AJ, Biesecker LG. A genotypic ascertainment approach to refute the association of MYO1A variants with non-syndromic deafness. Eur J Hum Genet. 2016;25(1):147-9. Epub 20161019. doi: 10.1038/ejhg.2016.140. PubMed PMID: 27759032; PubMed Central PMCID: 5159773.

426.   Bonnefond A, Yengo L, Philippe J, Dechaume A, Ezzidi I, Vaillant E, et al. Reassessment of the putative role of BLK-p.A71T loss-of-function mutation in

MODY and type 2 diabetes. Diabetologia. 2013;56(3):492-6. Epub 20121206. doi: 10.1007/s00125-012-2794-8. PubMed PMID: 23224494.

427.   Strande NT, Riggs ER, Buchanan AH, Ceyhan-Birsoy O, DiStefano M, Dwight SS, et al. Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. Am J Hum Genet. 2017;100(6):895-906. Epub 20170525. doi: 10.1016/j.ajhg.2017.04.015. PubMed PMID: 28552198; PubMed Central PMCID: 5473734.

428.   McGlaughon JL, Goldstein JL, Thaxton C, Hemphill SE, Berg JS. The progression of the ClinGen gene clinical validity classification over time. Hum Mutat. 2018;39(11):1494-504. doi: 10.1002/humu.23604. PubMed PMID: 30311372; PubMed Central PMCID: 6190678.

429.   Stark Z, Foulger RE, Williams E, Thompson BA, Patel C, Lunke S, et al. Scaling national and international improvement in virtual gene panel curation via a collaborative approach to discordance resolution. Am J Hum Genet. 2021;108(9):1551-7. Epub 20210729. doi: 10.1016/j.ajhg.2021.06.020. PubMed PMID: 34329581; PubMed Central PMCID: 8456155.

430.   Zhao WM, Seki A, Fang G. Cep55, a microtubule-bundling protein, associates with centralspindlin to control the midbody integrity and cell abscission during cytokinesis. Mol Biol Cell. 2006;17(9):3881-96. Epub 2006/06/21. doi: 10.1091/mbc.E06-01-0015. PubMed PMID: 16790497; PubMed Central PMCID: 1593165.

431.   Fabbro M, Zhou BB, Takahashi M, Sarcevic B, Lal P, Graham ME, et al. Cdk1/Erk2- and Plk1-dependent phosphorylation of a centrosome protein, Cep55, is required for its recruitment to midbody and cytokinesis. Dev Cell. 2005;9(4):477-88. doi: 10.1016/j.devcel.2005.09.003. PubMed PMID: 16198290.

432.   Xu ZY, Ma XS, Qi ST, Wang ZB, Guo L, Schatten H, et al. Cep55 regulates spindle organization and cell cycle progression in meiotic oocyte. Sci Rep. 2015;5:16978. Epub 2015/11/19. doi: 10.1038/srep16978. PubMed PMID: 26582107; PubMed Central PMCID: 4652202.

433.   Lee HH, Elia N, Ghirlando R, Lippincott-Schwartz J, Hurley JH. Midbody targeting of the ESCRT machinery by a noncanonical coiled coil in CEP55. Science. 2008;322(5901):576-80. doi: 10.1126/science.1162042. PubMed PMID: 18948538; PubMed Central PMCID: 2720046.

434.   Chen CH, Lu PJ, Chen YC, Fu SL, Wu KJ, Tsou AP, et al. FLJ10540-elicited cell transformation is through the activation of PI3-kinase/AKT pathway. Oncogene. 2007;26(29):4272-83. Epub 2007/01/22. doi: 10.1038/sj.onc.1210207. PubMed PMID: 17237822.

435.   Rashidieh B, Shohayeb B, Bain AL, Fortuna PRJ, Sinha D, Burgess A, et al. Cep55 regulation of PI3K/Akt signaling is required for neocortical development and ciliogenesis. PLoS Genet. 2021;17(10):e1009334. Epub 20211028. doi: 10.1371/journal.pgen.1009334. PubMed PMID: 34710087; PubMed Central PMCID: 8577787.

436.   Little JN, McNeely KC, Michel N, Bott CJ, Lettieri KS, Hecht MR, et al. Loss of Coiled-Coil Protein Cep55 Impairs Neural Stem Cell Abscission and Results in p53-Dependent Apoptosis in Developing Cortex. J Neurosci. 2021;41(15):3344-65. Epub 20210223. doi: 10.1523/JNEUROSCI.1955-20.2021. PubMed PMID: 33622776; PubMed Central PMCID: 8051691.

437.   Frosk P, Arts HH, Philippe J, Gunn CS, Brown EL, Chodirker B, et al. A truncating mutation in CEP55 is the likely cause of MARCH, a novel syndrome affecting neuronal mitosis. J Med Genet. 2017;54(7):490-501. Epub 2017/03/06. doi: 10.1136/jmedgenet-2016-104296. PubMed PMID: 28264986; PubMed Central PMCID: 5502313.

438.   Bondeson ML, Ericson K, Gudmundsson S, Ameur A, Pontén F, Wesström J, et al. A nonsense mutation in CEP55 defines a new locus for a Meckel-like syndrome, an autosomal recessive lethal fetal ciliopathy. Clin Genet. 2017. Epub 2017/03/14. doi: 10.1111/cge.13012. PubMed PMID: 28295209.

439.   Zhang YC, Bai YF, Yuan JF, Shen XL, Xu YL, Jian XX, et al. CEP55 promotes cilia disassembly through stabilizing Aurora A kinase. J Cell Biol. 2021;220(2). doi: 10.1083/jcb.202003149. PubMed PMID: 33475699; PubMed Central PMCID: 7829976.

440.   Tedeschi A, Almagro J, Renshaw MJ, Messal HA, Behrens A, Petronczki M. Cep55 promotes cytokinesis of neural progenitors but is dispensable for most mammalian cell divisions. Nat Commun. 2020;11(1):1746. Epub 20200408. doi: 10.1038/s41467-020-15359-w. PubMed PMID: 32269212; PubMed Central PMCID: 7142149.

441.   Curry CJ, Jensen K, Holland J, Miller L, Hall BD. The Potter sequence: a clinical analysis of 80 cases. Am J Med Genet. 1984;19(4):679-702. doi: 10.1002/ajmg.1320190408. PubMed PMID: 6393764.

442.   Simpson MA, Cross HE, Cross L, Helmuth M, Crosby AH. Lethal cystic kidney disease in Amish neonates associated with homozygous nonsense mutation of NPHP3. Am J Kidney Dis. 2009;53(5):790-5. Epub 2009/03/20. doi: 10.1053/j.ajkd.2008.12.026. PubMed PMID: 19303681.

443.   Bendon RW, Siddiqi T, de Courten-Myers G, Dignan P. Recurrent developmental anomalies: 1. Syndrome of hydranencephaly with renal aplastic dysplasia; 2. Polyvalvular developmental heart defect. Am J Med Genet Suppl. 1987;3:357-65. PubMed PMID: 3130870.

444.   Gschwendtner A, Mairinger T, Soelder E, Alge A, Kreczy A. Hydranencephaly with renal dysgenesis: a coincidental finding? Case report with review of the literature. Gynecol Obstet Invest. 1997;44(3):206-10. PubMed PMID: 9359650.

445.   Hamby WB, Krauss RF, Beswick WF. Hydranencephaly; clinical diagnosis; presentation of 7 cases. Pediatrics. 1950;6(3):371-83. PubMed PMID: 14780790.

446.   Strauss S, Bouzouki M, Goldfarb H, Uppal V, Costales F. Antenatal ultrasound diagnosis of an unusual case of hydranencephaly. J Clin Ultrasound. 1984;12(7):420-2. PubMed PMID: 6438176.

447.   Chu G, Miller W, Norton M, Kinney H, Genest D, Folkerth R. Hydranencephaly with binucleate neurons-renal dysplasia-syndactyly syndrome in three siblings   J Neuropathol Exp Neurol. 1998;54:483.

448.   Barisic I, Boban L, Loane M, Garne E, Wellesley D, Calzolari E, et al. Meckel-Gruber Syndrome: a population-based study on prevalence, prenatal diagnosis, clinical features, and survival in Europe. Eur J Hum Genet. 2015;23(6):746-52. Epub 2014/09/03. doi: 10.1038/ejhg.2014.174. PubMed PMID: 25182137; PubMed Central PMCID: 4795048.

449.   Barrie ES, Overwater E, van Haelst MM, Motazacker MM, Truxal KV, Crist E, et al. Expanding the spectrum of CEP55-associated disease to viable phenotypes. Am J Med Genet A. 2020;182(5):1201-8. Epub 20200225. doi: 10.1002/ajmg.a.61512. PubMed PMID: 32100459.

450.   Liu Y, Bankaitis VA. Phosphoinositide phosphatases in cell biology and disease. Prog Lipid Res. 2010;49(3):201-17. Epub 20100105. doi: 10.1016/j.plipres.2009.12.001. PubMed PMID: 20043944; PubMed Central PMCID: 2873057.

451.   Martin TF. Phosphoinositide lipids as signaling molecules: common themes for signal transduction, cytoskeletal regulation, and membrane trafficking. Annu Rev Cell Dev Biol. 1998;14:231-64. doi: 10.1146/annurev.cellbio.14.1.231. PubMed PMID: 9891784.

452.   Sasaki T, Takasuga S, Sasaki J, Kofuji S, Eguchi S, Yamazaki M, et al. Mammalian phosphoinositide kinases and phosphatases. Prog Lipid Res. 2009;48(6):307-43. Epub 20090704. doi: 10.1016/j.plipres.2009.06.001. PubMed PMID: 19580826.

453.   Attree O, Olivos IM, Okabe I, Bailey LC, Nelson DL, Lewis RA, et al. The Lowe's oculocerebrorenal syndrome gene encodes a protein highly homologous to inositol polyphosphate-5-phosphatase. Nature. 1992;358(6383):239-42. doi: 10.1038/358239a0. PubMed PMID: 1321346.

454.   Bielas SL, Silhavy JL, Brancati F, Kisseleva MV, Al-Gazali L, Sztriha L, et al. Mutations in INPP5E, encoding inositol polyphosphate-5-phosphatase E, link phosphatidyl inositol signaling to the ciliopathies. Nat Genet. 2009;41(9):1032-6. Epub 20090809. doi: 10.1038/ng.423. PubMed PMID: 19668216; PubMed Central PMCID: 2746682.

455.   Wiessner M, Roos A, Munn CJ, Viswanathan R, Whyte T, Cox D, et al. Mutations in INPP5K, Encoding a Phosphoinositide 5-Phosphatase, Cause Congenital Muscular Dystrophy with Cataracts and Mild Cognitive Impairment. Am J Hum Genet. 2017;100(3):523-36. Epub 20170209. doi: 10.1016/j.ajhg.2017.01.024. PubMed PMID: 28190456; PubMed Central PMCID: 5339217.

456. Hardies K, Cai Y, Jardel C, Jansen AC, Cao M, May P, et al. Loss of SYNJ1 dual phosphatase activity leads to early onset refractory seizures and progressive neurological decline. Brain. 2016;139(Pt 9):2420-30. Epub 20160719. doi: 10.1093/brain/aww180. PubMed PMID: 27435091; PubMed Central PMCID: 4995362.

457. Norris FA, Majerus PW. Hydrolysis of phosphatidylinositol 3,4-bisphosphate by inositol polyphosphate 4-phosphatase isolated by affinity elution chromatography. J Biol Chem. 1994;269(12):8716-20. PubMed PMID: 8132601.

458. Ivetac I, Gurung R, Hakim S, Horan KA, Sheffield DA, Binge LC, et al. Regulation of PI(3)K/Akt signalling and cellular transformation by inositol polyphosphate 4-phosphatase-1. EMBO Rep. 2009;10(5):487-93. Epub 20090327. doi: 10.1038/embor.2009.28. PubMed PMID: 19325558; PubMed Central PMCID: 2680870.

459. Li H, Marshall AJ. Phosphatidylinositol (3,4) bisphosphate-specific phosphatases and effector proteins: A distinct branch of PI3K signaling. Cell Signal. 2015;27(9):1789-98. Epub 20150527. doi: 10.1016/j.cellsig.2015.05.013. PubMed PMID: 26022180.

460. Nystuen A, Legare ME, Shultz LD, Frankel WN. A null mutation in inositol polyphosphate 4-phosphatase type I causes selective neuronal loss in weeble mutant mice. Neuron. 2001;32(2):203-12. doi: 10.1016/s0896-6273(01)00468-8. PubMed PMID: 11683991.

461. Wang L, Wang Y, Duan C, Yang Q. Inositol phosphatase INPP4A inhibits the apoptosis of. Int J Clin Exp Pathol. 2018;11(4):1999-2007. Epub 20180401. PubMed PMID: 31938306; PubMed Central PMCID: 6958220.

462. Sasaki J, Kofuji S, Itoh R, Momiyama T, Takayama K, Murakami H, et al. The PtdIns(3,4)P(2) phosphatase INPP4A is a suppressor of excitotoxic neuronal death. Nature. 2010;465(7297):497-501. Epub 20100512. doi: 10.1038/nature09023. PubMed PMID: 20463662.

463. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011;478(7367):57-63. Epub 2011/09/21. doi: 10.1038/nature10423. PubMed PMID: 21937992.

464. Sheffer R, Bennett-Back O, Yaacov B, Edvardson S, Gomori M, Werner M, et al. Hindbrain malformation and myoclonic seizures associated with a deleterious mutation in the INPP4A gene. Neurogenetics. 2015;16(1):23-6. Epub 20141022. doi: 10.1007/s10048-014-0428-7. PubMed PMID: 25338135.

465. Olson HE, Kelly M, LaCoursiere CM, Pinsky R, Tambunan D, Shain C, et al. Genetics and genotype-phenotype correlations in early onset epileptic encephalopathy with burst suppression. Ann Neurol. 2017;81(3):419-29. Epub 20170214. doi: 10.1002/ana.24883. PubMed PMID: 28133863; PubMed Central PMCID: 5366084.

466. Banihashemi S, Tahmasebi-Birgani M, Mohammadiasl J, Hajjari M. Whole exome sequencing identified a novel nonsense INPP4A mutation in a family with intellectual disability. Eur J Med Genet. 2020;63(4):103846. Epub 20200121. doi: 10.1016/j.ejmg.2020.103846. PubMed PMID: 31978615.

467. Lindeboom RG, Supek F, Lehner B. The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat Genet. 2016;48(10):1112-8. Epub 20160912. doi: 10.1038/ng.3664. PubMed PMID: 27618451; PubMed Central PMCID: 5045715.

468. Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem Sci. 1998;23(6):198-9. doi: 10.1016/s0968-0004(98)01208-0. PubMed PMID: 9644970.

469. Chaudhuri R, Khanna K, Koundinya D, Pattnaik B, Vatsa D, Agrawal A, et al. Novel nuclear translocation of inositol polyphosphate 4-phosphatase is associated with cell cycle, proliferation and survival. Biochim Biophys Acta Mol Cell Res. 2018. Epub 20180730. doi: 10.1016/j.bbamcr.2018.07.013. PubMed PMID: 30071275.

470. Sachs AJ, David SA, Haider NB, Nystuen AM. Patterned neuroprotection in the Inpp4a(wbl) mutant mouse cerebellum correlates with the expression of Eaat4. PLoS One. 2009;4(12):e8270. Epub 20091214. doi: 10.1371/journal.pone.0008270. PubMed PMID: 20011524; PubMed Central PMCID: 2788419.

471. Held PK, Rice GM, Kuhl A, Drilias N, Baker M, Deline J, et al. Newborn Screening for Inherited Metabolic Disorders: Early Identification and Long-Term Care for Patients in the Plain Community, Wisconsin, 2011-2017. Public Health Rep. 2019;134(2_suppl):58S-63S. doi: 10.1177/0033354919878425. PubMed PMID: 31682555; PubMed Central PMCID: 6832032.

472. Kuhl A, van Calcar S, Baker M, Seroogy CM, Rice G, Scott Schwoerer J. Development of carrier testing for common inborn errors of metabolism in the Wisconsin Plain population. Genet Med. 2017;19(3):352-6. Epub 20160811. doi: 10.1038/gim.2016.104. PubMed PMID: 27513192.

473. Schwoerer JS, Drilias N, Kuhl A, Mochal S, Baker M. Genotypes of patients with phenylalanine hydroxylase deficiency in the Wisconsin Amish. Mol Genet Metab Rep. 2018;15:75-7. Epub 20180308. doi: 10.1016/j.ymgmr.2018.02.005. PubMed PMID: 29560316; PubMed Central PMCID: 5857495.

474. Scott Schwoerer J, Clowes Candadai S, Held PK. Long-term outcomes in Amish patients diagnosed with propionic acidemia. Mol Genet Metab Rep. 2018;16:36-8. Epub 20180622. doi: 10.1016/j.ymgmr.2018.05.004. PubMed PMID: 30013935; PubMed Central PMCID: 6019757.

475. Singh NA, Pappas C, Dahle EJ, Claes LR, Pruess TH, De Jonghe P, et al. A role of SCN9A in human epilepsies, as a cause of febrile seizures and as a potential modifier of Dravet syndrome. PLoS Genet. 2009;5(9):e1000649. Epub

20090918. doi: 10.1371/journal.pgen.1000649. PubMed PMID: 19763161; PubMed Central PMCID: 2730533.

476. Ameur A, Kloosterman WP, Hestand MS. Single-Molecule Sequencing: Towards Clinical Applications. Trends Biotechnol. 2019;37(1):72-85. Epub 20180813. doi: 10.1016/j.tibtech.2018.07.013. PubMed PMID: 30115375.

477. Watson CM, Jackson L, Crinnion LA, Bonthron DT, Sheridan E. Long-read sequencing to resolve the parent of origin of a de novo pathogenic. J Med Genet. 2022. Epub 20220412. doi: 10.1136/jmedgenet-2021-108314. PubMed PMID: 35414530.

478. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-308. Epub 2013/10/24. doi: 10.1038/nprot.2013.143. PubMed PMID: 24157548; PubMed Central PMCID: 3969860.

479. Skarnes WC, Rosen B, West AP, Koutsourakis M, Bushell W, Iyer V, et al. A conditional knockout resource for the genome-wide study of mouse gene function. Nature. 2011;474(7351):337-42. Epub 2011/06/15. doi: 10.1038/nature10163. PubMed PMID: 21677750; PubMed Central PMCID: 3572410.

480. Collins SC, Mikhaleva A, Vrcelj K, Vancollie VE, Wagner C, Demeure N, et al. Large-scale neuroanatomical study uncovers 198 gene associations in mouse brain morphogenesis. Nat Commun. 2019;10(1):3465. Epub 2019/08/01. doi: 10.1038/s41467-019-11431-2. PubMed PMID: 31371714; PubMed Central PMCID: 6671969.

481. Sunkin SM, Ng L, Lau C, Dolbeare T, Gilbert TL, Thompson CL, et al. Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Res. 2013;41(Database issue):D996-D1008. Epub 2012/11/28. doi: 10.1093/nar/gks1042. PubMed PMID: 23193282; PubMed Central PMCID: 3531093.

482. Collins SC, Wagner C, Gagliardi L, Kretz PF, Fischer MC, Kessler P, et al. A Method for Parasagittal Sectioning for Neuroanatomical Quantification of Brain Structures in the Adult Mouse. Curr Protoc Mouse Biol. 2018;8(3):e48. Epub 2018/06/26. doi: 10.1002/cpmo.48. PubMed PMID: 29944194.

483. Gavin AC, Bösche M, Krause R, Grandi P, Marzioch M, Bauer A, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature. 2002;415(6868):141-7. doi: 10.1038/415141a. PubMed PMID: 11805826.

484. Kümmel D, Müller JJ, Roske Y, Henke N, Heinemann U. Structure of the Bet3-Tpc6B core of TRAPP: two Tpc6 paralogs form trimeric complexes with Bet3 and Mum2. J Mol Biol. 2006;361(1):22-32. Epub 20060621. doi: 10.1016/j.jmb.2006.06.012. PubMed PMID: 16828797.

485. Sacher M, Jiang Y, Barrowman J, Scarpa A, Burston J, Zhang L, et al. TRAPP, a highly conserved novel complex on the cis-Golgi that mediates vesicle

docking and fusion. EMBO J. 1998;17(9):2494-503. doi: 10.1093/emboj/17.9.2494. PubMed PMID: 9564032; PubMed Central PMCID: 1170591.

486.   Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. Nucleic Acids Res. 2021;49(D1):D412-D9. doi: 10.1093/nar/gkaa913. PubMed PMID: 33125078; PubMed Central PMCID: 7779014.

487.   Kim YG, Sohn EJ, Seo J, Lee KJ, Lee HS, Hwang I, et al. Crystal structure of bet3 reveals a novel mechanism for Golgi localization of tethering factor TRAPP. Nat Struct Mol Biol. 2005;12(1):38-45. Epub 20041219. doi: 10.1038/nsmb871. PubMed PMID: 15608655.

488.   Zong M, Wu XG, Chan CW, Choi MY, Chan HC, Tanner JA, et al. The adaptor function of TRAPPC2 in mammalian TRAPPs explains TRAPPC2-associated SEDT and TRAPPC9-associated congenital intellectual disability. PLoS One. 2011;6(8):e23350. Epub 2011/08/15. doi: 10.1371/journal.pone.0023350. PubMed PMID: 21858081; PubMed Central PMCID: 3156116.

489.   Choi C, Davey M, Schluter C, Pandher P, Fang Y, Foster LJ, et al. Organization and assembly of the TRAPPII complex. Traffic. 2011;12(6):715-25. Epub 20110401. doi: 10.1111/j.1600-0854.2011.01181.x. PubMed PMID: 21453443.

490.   Lamb CA, Nühlen S, Judith D, Frith D, Snijders AP, Behrends C, et al. TBC1D14 regulates autophagy via the TRAPP complex and ATG9 traffic. EMBO J. 2016;35(3):281-301. Epub 20151228. doi: 10.15252/embj.201592695. PubMed PMID: 26711178; PubMed Central PMCID: 4741301.

491.   Schou KB, Morthorst SK, Christensen ST, Pedersen LB. Identification of conserved, centrosome-targeting ASH domains in TRAPPII complex subunits and TRAPPC8. Cilia. 2014;3:6. Epub 20140618. doi: 10.1186/2046-2530-3-6. PubMed PMID: 25018876; PubMed Central PMCID: 4094338.

492.   Turnbull AP, Kümmel D, Prinz B, Holz C, Schultchen J, Lang C, et al. Structure of palmitoylated BET3: insights into TRAPP complex assembly and membrane localization. EMBO J. 2005;24(5):875-84. Epub 20050203. doi: 10.1038/sj.emboj.7600565. PubMed PMID: 15692564; PubMed Central PMCID: 554119.

493.   Nachury MV, Loktev AV, Zhang Q, Westlake CJ, Peränen J, Merdes A, et al. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. Cell. 2007;129(6):1201-13. doi: 10.1016/j.cell.2007.03.053. PubMed PMID: 17574030.

494.   Al-Dosari MS, Al-Shammari M, Shaheen R, Faqeih E, Alghofely MA, Boukai A, et al. 3M syndrome: an easily recognizable yet underdiagnosed cause of proportionate short stature. J Pediatr. 2012;161(1):139-45.e1. Epub 20120209. doi: 10.1016/j.jpeds.2011.12.051. PubMed PMID: 22325252.

495. Wendler F, Gillingham AK, Sinka R, Rosa-Ferreira C, Gordon DE, Franch-Marro X, et al. A genome-wide RNA interference screen identifies two novel components of the metazoan secretory pathway. EMBO J. 2010;29(2):304-14. Epub 20091126. doi: 10.1038/emboj.2009.350. PubMed PMID: 19942856; PubMed Central PMCID: 2824459.

496. Kim JJ, Lipatova Z, Segev N. TRAPP Complexes in Secretion and Autophagy. Front Cell Dev Biol. 2016;4:20. Epub 2016/03/30. doi: 10.3389/fcell.2016.00020. PubMed PMID: 27066478; PubMed Central PMCID: 4811894.

497. Stanga D, Zhao Q, Milev MP, Saint-Dic D, Jimenez-Mallebrera C, Sacher M. TRAPPC11 functions in autophagy by recruiting ATG2B-WIPI4/WDR45 to preautophagosomal membranes. Traffic. 2019;20(5):325-45. Epub 20190409. doi: 10.1111/tra.12640. PubMed PMID: 30843302.

498. Milev MP, Hasaj B, Saint-Dic D, Snounou S, Zhao Q, Sacher M. TRAMM/TrappC12 plays a role in chromosome congression, kinetochore stability, and CENP-E recruitment. J Cell Biol. 2015;209(2):221-34. doi: 10.1083/jcb.201501090. PubMed PMID: 25918224; PubMed Central PMCID: 4411272.

499. Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. Nature. 2010;466(7302):68-76. Epub 20100620. doi: 10.1038/nature09204. PubMed PMID: 20562859; PubMed Central PMCID: 2901998.

500. Ramírez-Peinado S, Ignashkova TI, van Raam BJ, Baumann J, Sennott EL, Gendarme M, et al. TRAPPC13 modulates autophagy and the response to Golgi stress. J Cell Sci. 2017;130(14):2251-65. Epub 20170523. doi: 10.1242/jcs.199521. PubMed PMID: 28536105; PubMed Central PMCID: 6518214.

501. Caulfield M, Davies J, Dennys M, Elbahy L, Fowler T, Hill S. *The National Genomics Research and Healthcare Knowledgebase v7, Genomics England*. 2020.

502. Gandolfo LC, Bahlo M, Speed TP. Dating rare mutations from small samples with dense marker data. Genetics. 2014;197(4):1315-27. Epub 2014/05/30. doi: 10.1534/genetics.114.164616. PubMed PMID: 24879464; PubMed Central PMCID: 4125402.

503. Testa G, Schaft J, van der Hoeven F, Glaser S, Anastassiadis K, Zhang Y, et al. A reliable lacZ expression reporter cassette for multipurpose, knockout-first alleles. Genesis. 2004;38(3):151-8. doi: 10.1002/gene.20012. PubMed PMID: 15048813.

504. Dias C, Estruch SB, Graham SA, McRae J, Sawiak SJ, Hurst JA, et al. BCL11A Haploinsufficiency Causes an Intellectual Disability Syndrome and Dysregulates Transcription. Am J Hum Genet. 2016;99(2):253-74. Epub 20160721. doi: 10.1016/j.ajhg.2016.05.030. PubMed PMID: 27453576; PubMed Central PMCID: 4974071.

505. Morton AJ, Skillings E, Bussey TJ, Saksida LM. Measuring cognitive deficits in disabled mice using an automated interactive touchscreen system. Nat Methods. 2006;3(10):767. doi: 10.1038/nmeth1006-767. PubMed PMID: 16990806.

506. Horner AE, Heath CJ, Hvoslef-Eide M, Kent BA, Kim CH, Nilsson SR, et al. The touchscreen operant platform for testing learning and memory in rats and mice. Nat Protoc. 2013;8(10):1961-84. Epub 20130919. doi: 10.1038/nprot.2013.122. PubMed PMID: 24051959; PubMed Central PMCID: 3914026.

507. Bubser M, Bridges TM, Dencker D, Gould RW, Grannan M, Noetzel MJ, et al. Selective activation of M4 muscarinic acetylcholine receptors reverses MK-801-induced behavioral impairments and enhances associative learning in rodents. ACS Chem Neurosci. 2014;5(10):920-42. Epub 20140819. doi: 10.1021/cn500128b. PubMed PMID: 25137629; PubMed Central PMCID: 4324418.

508. Brigman JL, Daut RA, Wright T, Gunduz-Cinar O, Graybeal C, Davis MI, et al. GluN2B in corticostriatal circuits governs choice learning and choice shifting. Nat Neurosci. 2013;16(8):1101-10. Epub 20130707. doi: 10.1038/nn.3457. PubMed PMID: 23831965; PubMed Central PMCID: 3725191.

509. Harrison FE, Reiserer RS, Tomarken AJ, McDonald MP. Spatial and nonspatial escape strategies in the Barnes maze. Learn Mem. 2006;13(6):809-19. Epub 20061113. doi: 10.1101/lm.334306. PubMed PMID: 17101874; PubMed Central PMCID: 1783636.

510. Harrison FE, Hosseini AH, McDonald MP. Endogenous anxiety and stress responses in water maze and Barnes maze spatial memory tasks. Behav Brain Res. 2009;198(1):247-51. Epub 20081018. doi: 10.1016/j.bbr.2008.10.015. PubMed PMID: 18996418; PubMed Central PMCID: 2663577.

511. Sawiak SJ, Wood NI, Williams GB, Morton AJ, Carpenter TA. Voxel-based morphometry with templates and validation in a mouse model of Huntington's disease. Magn Reson Imaging. 2013;31(9):1522-31. Epub 20130706. doi: 10.1016/j.mri.2013.06.001. PubMed PMID: 23835187; PubMed Central PMCID: 3919157.

512. Ashburner J, Friston KJ. Voxel-based morphometry--the methods. Neuroimage. 2000;11(6 Pt 1):805-21. doi: 10.1006/nimg.2000.0582. PubMed PMID: 10860804.

513. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21. Epub 20121025. doi: 10.1093/bioinformatics/bts635. PubMed PMID: 23104886; PubMed Central PMCID: 3530905.

514. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923-30. Epub 20131113. doi: 10.1093/bioinformatics/btt656. PubMed PMID: 24227677.

515.   Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PubMed Central PMCID: 4302049.

516.   Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet. 2007;3(9):1724-35. Epub 20070801. doi: 10.1371/journal.pgen.0030161. PubMed PMID: 17907809; PubMed Central PMCID: 1994707.

517.   Kilpinen H, Goncalves A, Leha A, Afzal V, Alasoo K, Ashford S, et al. Common genetic variation drives molecular heterogeneity in human iPSCs. Nature. 2017;546(7658):370-5. Epub 20170510. doi: 10.1038/nature22403. PubMed PMID: 28489815; PubMed Central PMCID: 5524171.

518.   Subramanian A, Narayan R, Corsello SM, Peck DD, Natoli TE, Lu X, et al. A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell. 2017;171(6):1437-52.e17. doi: 10.1016/j.cell.2017.10.049. PubMed PMID: 29195078; PubMed Central PMCID: 5990023.

519.   Chudasama Y, Robbins TW. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. J Neurosci. 2003;23(25):8771-80. PubMed PMID: 14507977; PubMed Central PMCID: 6740430.

520.   Izquierdo A, Darling C, Manos N, Pozos H, Kim C, Ostrander S, et al. Basolateral amygdala lesions facilitate reward choices after negative feedback in rats. J Neurosci. 2013;33(9):4105-9. doi: 10.1523/JNEUROSCI.4942-12.2013. PubMed PMID: 23447618; PubMed Central PMCID: PMCPMC3606920.

521.   Kwon CH, Zhu X, Zhang J, Baker SJ. mTor is required for hypertrophy of Pten-deficient neuronal soma in vivo. Proc Natl Acad Sci U S A. 2003;100(22):12923-8. Epub 20031008. doi: 10.1073/pnas.2132711100. PubMed PMID: 14534328; PubMed Central PMCID: 240720.

522.   Klintsova AY, Hamilton GF, Boschen KE. Long-term consequences of developmental alcohol exposure on brain structure and function: therapeutic benefits of physical activity. Brain Sci. 2012;3(1):1-38. Epub 20121221. doi: 10.3390/brainsci3010001. PubMed PMID: 24961305; PubMed Central PMCID: 4061829.

523.   Semple BD, Blomgren K, Gimlin K, Ferriero DM, Noble-Haeusslein LJ. Brain development in rodents and humans: Identifying benchmarks of maturation and vulnerability to injury across species. Prog Neurobiol. 2013;106-107:1-16. Epub 20130411. doi: 10.1016/j.pneurobio.2013.04.001. PubMed PMID: 23583307; PubMed Central PMCID: 3737272.

524.   Rozengurt E, Soares HP, Sinnet-Smith J. Suppression of feedback loops mediated by PI3K/mTOR induces multiple overactivation of compensatory pathways: an unintended consequence leading to drug resistance. Mol Cancer

Ther. 2014;13(11):2477-88. Epub 20141016. doi: 10.1158/1535-7163.MCT-14-0330. PubMed PMID: 25323681; PubMed Central PMCID: 4222988.

525.   Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):W191-W8. doi: 10.1093/nar/gkz369. PubMed PMID: 31066453; PubMed Central PMCID: 6602461.

526.   Cardenas ME, Cutler NS, Lorenz MC, Di Como CJ, Heitman J. The TOR signaling cascade regulates gene expression in response to nutrients. Genes Dev. 1999;13(24):3271-9. doi: 10.1101/gad.13.24.3271. PubMed PMID: 10617575; PubMed Central PMCID: 317202.

527.   Mayer C, Grummt I. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. Oncogene. 2006;25(48):6384-91. doi: 10.1038/sj.onc.1209883. PubMed PMID: 17041624.

528.   Wang X, Proud CG. The mTOR pathway in the control of protein synthesis. Physiology (Bethesda). 2006;21:362-9. doi: 10.1152/physiol.00024.2006. PubMed PMID: 16990457.

529.   Huang J, Manning BD. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. Biochem J. 2008;412(2):179-90. doi: 10.1042/BJ20080281. PubMed PMID: 18466115; PubMed Central PMCID: 2735030.

530.   Winden KD, Sundberg M, Yang C, Wafa SMA, Dwyer S, Chen PF, et al. Biallelic mutations in TSC2 lead to abnormalities associated with cortical tubers in human ipsc-derived neurons. J Neurosci. 2019;39(47):9294-305. Epub 20191007. doi: 10.1523/JNEUROSCI.0642-19.2019. PubMed PMID: 31591157; PubMed Central PMCID: 6867816.

531.   Liu Q, Xu C, Kirubakaran S, Zhang X, Hur W, Liu Y, et al. Characterization of Torin2, an ATP-competitive inhibitor of mTOR, ATM, and ATR. Cancer Res. 2013;73(8):2574-86. Epub 20130222. doi: 10.1158/0008-5472.CAN-12-1702. PubMed PMID: 23436801; PubMed Central PMCID: 3760004.

532.   Wang J, Lin ZJ, Liu L, Xu HQ, Shi YW, Yi YH, et al. Epilepsy-associated genes. Seizure. 2017;44:11-20. Epub 20161206. doi: 10.1016/j.seizure.2016.11.030. PubMed PMID: 28007376.

533.   Thormann A, Halachev M, McLaren W, Moore DJ, Svinti V, Campbell A, et al. Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. Nat Commun. 2019;10(1):2373. Epub 20190530. doi: 10.1038/s41467-019-10016-3. PubMed PMID: 31147538; PubMed Central PMCID: 6542828.

534. Rogawski MA. KCNQ2/KCNQ3 K+ channels and the molecular pathogenesis of epilepsy: implications for therapy. Trends Neurosci.

2000;23(9):393-8. doi: 10.1016/s0166-2236(00)01629-5. PubMed PMID: 10941184.

535. Maljevic S, Vejzovic S, Bernhard MK, Bertsche A, Weise S, Döcker M, et al. Novel *KCNQ3* Mutation in a Large Family with Benign Familial Neonatal Epilepsy: A Rare Cause of Neonatal Seizures. Mol Syndromol. 2016;7(4):189-96. Epub 20160707. doi: 10.1159/000447461. PubMed PMID: 27781029; PubMed Central PMCID: 5073621.

536. Hansen J, Snow C, Tuttle E, Ghoneim DH, Yang CS, Spencer A, et al. De novo mutations in SIK1 cause a spectrum of developmental epilepsies. Am J Hum Genet. 2015;96(4):682-90. doi: 10.1016/j.ajhg.2015.02.013. PubMed PMID: 25839329; PubMed Central PMCID: 4385182.

537. Pröschel C, Hansen JN, Ali A, Tuttle E, Lacagnina M, Buscaglia G, et al. Epilepsy-causing sequence variations in SIK1 disrupt synaptic activity response gene expression and affect neuronal morphology. Eur J Hum Genet. 2017;25(2):216-21. Epub 20161214. doi: 10.1038/ejhg.2016.145. PubMed PMID: 27966542; PubMed Central PMCID: 5255945.

538. Andrews MG, Subramanian L, Kriegstein AR. mTOR signaling regulates the morphology and migration of outer radial glia in developing human cortex. Elife. 2020;9. Epub 20200902. doi: 10.7554/eLife.58737. PubMed PMID: 32876565; PubMed Central PMCID: 7467727.

539. Knobloch M, Braun SM, Zurkirchen L, von Schoultz C, Zamboni N, Araúzo-Bravo MJ, et al. Metabolic control of adult neural stem cell activity by Fasn-dependent lipogenesis. Nature. 2013;493(7431):226-30. Epub 20121202. doi: 10.1038/nature11689. PubMed PMID: 23201681; PubMed Central PMCID: 3587167.

540. Knobloch M, von Schoultz C, Zurkirchen L, Braun SM, Vidmar M, Jessberger S. SPOT14-positive neural stem/progenitor cells in the hippocampus respond dynamically to neurogenic regulators. Stem Cell Reports. 2014;3(5):735-42. Epub 20140926. doi: 10.1016/j.stemcr.2014.08.013. PubMed PMID: 25418721; PubMed Central PMCID: 4235138.

541. Shin J, Berg DA, Zhu Y, Shin JY, Song J, Bonaguidi MA, et al. Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. Cell Stem Cell. 2015;17(3):360-72. Epub 20150820. doi: 10.1016/j.stem.2015.07.013. PubMed PMID: 26299571; PubMed Central PMCID: 8638014.

542. Hochgerner H, Zeisel A, Lönnerberg P, Linnarsson S. Conserved properties of dentate gyrus neurogenesis across postnatal development revealed by single-cell RNA sequencing. Nat Neurosci. 2018;21(2):290-9. Epub 20180115. doi: 10.1038/s41593-017-0056-2. PubMed PMID: 29335606.

543. Di Bella DJ, Habibi E, Stickels RR, Scalia G, Brown J, Yadollahpour P, et al. Molecular logic of cellular diversification in the mouse cerebral cortex. Nature. 2021;595(7868):554-9. Epub 20210623. doi: 10.1038/s41586-021-03670-5. PubMed PMID: 34163074.

544.	Switon K, Kotulska K, Janusz-Kaminska A, Zmorzynska J, Jaworski J. Molecular neurobiology of mTOR. Neuroscience. 2017;341:112-53. Epub 20161123. doi: 10.1016/j.neuroscience.2016.11.017. PubMed PMID: 27889578.

545.	Kaplanis J, Samocha KE, Wiel L, Zhang Z, Arvai KJ, Eberhardt RY, et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. Nature. 2020;586(7831):757-62. Epub 20201014. doi: 10.1038/s41586-020-2832-5. PubMed PMID: 33057194; PubMed Central PMCID: 7116826.