



Neurodevelopmental disorders: **a next generation**

Margot R. F. Reijnders

Neurodevelopmental disorders: a next generation

Margot R.F. Reijnders

The work presented in this thesis was carried out within the Donders Institute for Brain, Cognition and Behaviour, at the Department of Human Genetics of the Radboudumc in Nijmegen, The Netherlands.

ISBN: 978-94-9289-699-5

Cover design:
Frankemargrete.nl

Lay-out:
Michiel van Zeijl, Franke Hoogerkamp and Margot Reijnders

Printed by:
Ipskamp printing

© 2018, Margot Reijnders, Arnhem, The Netherlands

All rights reserved. No part of this publication may be reproduced, stored in retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission.

Supplemental data of this thesis are available at:



Neurodevelopmental disorders: a next generation

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op vrijdag 11 januari 2019
om 10.30 uur precies

door

Margot Rosa Frieda Reijnders
geboren op 27 juli 1989
te Roosendaal en Nispen

Promotor:

Prof. dr. H.G. Brunner

Copromotoren:

Dr. T. Kleefstra

Dr. L.E.L.M. Vissers

Manuscriptcommissie:

Prof. A. Cambi

Prof. M.A.A.P. Willemsen

Prof. B.L. Loeys (Universiteit Antwerpen, België)

Neurodevelopmental disorders: a next generation

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of the Rector Magnificus prof. dr. J.H.J.M. van Krieken,
according to the decision of the Council of Deans
to be defended in public on Friday, January 11, 2019
at 10.30 hours

by

Margot Rosa Frieda Reijnders
Born on July 27, 1989
in Roosendaal en Nispen (The Netherlands)

Supervisor:

Prof. dr. H.G. Brunner

Co-supervisors:

Dr. T. Kleefstra

Dr. L.E.L.M. Vissers

Doctoral Thesis Committee:

Prof. A. Cambi

Prof. M.A.A.P. Willemsen

Prof. B.L. Loeys (Antwerp University, Belgium)

Table of contents

1	General introduction	10
	1.1 Neurodevelopmental disorders	12
	1.2 Establishing a genetic diagnosis in patients with neurodevelopmental disorders	15
	1.3 Whole Exome Sequencing: a complex diagnostic work-flow	16
	1.4 Diagnostic yield of Whole Exome Sequencing	21
2	Scope and outline of this thesis	22
<hr/>		
3	Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability <i>Nature Neuroscience (2016) 19, 1194-1196</i>	30
4	<i>De novo</i> mutations in <i>SON</i> disrupt RNA splicing of genes essential for brain development and metabolism, causing an intellectual disability syndrome <i>American Journal of Human Genetics (2016) 99, 711-719</i>	46
5	<i>De novo</i> and inherited mutations in <i>TLK2</i> : identification, clinical delineation and genotype-phenotype evaluation of a novel neurodevelopmental disorder <i>American Journal of Human Genetics (2018) 102, 1195-1203</i>	58
6	<i>RAC1</i> mutations in developmental disorders with diverse phenotypes <i>American Journal of Human Genetics (2017) 101, 466-477</i>	72
7	Variation in a range of mTOR-related genes associates with intracranial volume and intellectual disability <i>Nature Communications (2017) 8, 1052</i>	88

8	Recurrent <i>de novo</i> heterozygous mutations disturbing the GTP/GDP binding pocket of <i>RAB11B</i> cause intellectual disability and a distinctive brain phenotype <i>American Journal of Human Genetics</i> (2017) 101, 824-832	110
9	<i>De novo</i> loss-of-function mutations in <i>USP9X</i> cause a female-specific recognizable syndrome with developmental delay and congenital malformations <i>American Journal of Human Genetics</i> (2016) 98, 373-381	124
10	<i>De novo</i> loss-of function mutations in <i>WAC</i> cause a recognizable intellectual disability syndrome and are associated with learning deficits in <i>Drosophila</i> <i>European Journal of Human Genetics</i> (2016) 24, 1145-53	136
11	PURA syndrome: clinical delineation and genotype-phenotype study in 32 patients with review of published literature <i>Journal of Medical Genetics</i> (2017) 55, 104-113	152
<hr/>		
12	General discussion and future perspective	170
	12.1 Neurodevelopmental disorders: a next generation	172
	12.1.1 Next generation NDDs & genetics: discovery of rare NDDs	172
	12.1.2 Next generation NDDs & clinic: confirmation and delineation by matchmaking	178
	12.1.3 Next generation NDDs & biology: unraveling underlying biological mechanism	184
	12.2 Neurodevelopmental disorders: what's next?	190
	12.2.1 Diagnostic and research strategies for undiagnosed NDD patients	190
	12.2.2 Improved patient care for next generation NDDs	197
<hr/>		

A

Appendices	200
A. Reference list	202
B. Nederlandse samenvatting	236
C. Dankwoord / Acknowledgements	240
D. Curriculum Vitae	248
E. List of publications	250



1

General introduction

1

1.1 Neurodevelopmental disorders

Definition

Neurodevelopmental disorders (NDDs) are a group of conditions characterized by developmental deficits that produce impairments of personal, social, academic, or occupational functioning. These disorders typically become manifest before a child enters elementary school (age 6 years).¹ Six main subtypes of NDDs can be distinguished (Table 1.1):

- 1. Intellectual disability (ID) or global developmental delay (DD) if age is below 5 years
- 2. Communication disorder (CD)
- 3. Autism spectrum disorder (ASD)
- 4. Attention-deficit hyperactivity disorder (ADHD)
- 5. Specific learning disorder (SLD)
- 6. Motor disorder (MD)

Different subtypes of NDDs can be present in a single patient, and the severity of the disorders can vary between patients and over time. Subtypes of NDDs have also been defined in guidelines published by the World Health Organization (ICD-10, released in 2016). In general, the section ‘Mental and behavioral disorders’ contains definitions for such subtypes, but is not restricted to NDDs, as it also includes other psychiatric diagnosis, such as schizophrenia, mood disorders and eating disorders.²

Box 1: DSM and Autism Spectrum Disorders

The Diagnostic and Statistical Manual of Mental Disorders (DSM) is a handbook used by health care professionals as authoritative guide to the diagnosis of mental disorders. Since the introduction in 1952, it has been revised six times. The last revision was in 2013, with the publication of the DSM-5 criteria. These criteria combined what previously were four separate disorders into one condition:



In this thesis, the term ‘Autism Spectrum Disorder’ is not always used to describe a psychiatric diagnosis for patients. Several patients received their psychiatric diagnosis before the introduction of the new DSM criteria in 2013 or in the transition period from DSM-IV to DSM-5. For this reason, diagnostic terms such as PDD-NOS and Autism (DSM-IV) are used several times, instead of ‘Autism Spectrum Disorders’ (DMS-5). Additionally, the severity scoring of autism spectrum disorders, introduced in DSM-5, is often not available.

Table 1.1 Diagnostic criteria of different subtypes of neurodevelopmental disorders*

Disorder	Diagnostic criteria**	Diagnostic terms
Intellectual disability (ID)	(1) Deficits in intellectual functioning, confirmed by both clinical assessment and individualized, standardized intelligence testing. (2) Deficits in adaptive functioning that result in failure to meet developmental and sociocultural standards for personal independence and social responsibility.	Intellectual disability; Global developmental delay
Communication disorder (CD)	(1) Persistent difficulties in the acquisition and use of language across modalities due to deficits in comprehension or production. (2) Language abilities are substantially and quantifiably below those expected for age, resulting in functional limitations. (3) The difficulties are not attributable to other disorders.	Language disorder; Speech sound disorder; Childhood-onset fluency disorder; Social communication disorder
Autism spectrum disorder (ASD)	(1) Persistent deficits in social communication and social interaction across multiple contexts. (2) Restricted, repetitive patterns of behavioral interest, or activities. (3) Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning. (4) Disturbances are not better explained by ID or global developmental delay.	Autism spectrum disorder
Attention-deficit hyperactivity disorder (ADHD)	(1) Persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning of development. (2) Several inattentive or hyperactive-impulsive symptoms are present in two or more settings. (3) There is clear evidence that the symptoms interfere with, or reduce the quality of, social, academic, or occupational functioning. (4) The symptoms do not occur exclusively during the course of a psychotic disorder and are not better explained by another mental disorder.	Attention deficit/hyperactivity disorder; Other specified attention deficit/hyperactivity disorder
Specific learning disorder (SLD)	(1) Difficulties with learning and using academic skills that have persisted for at least six months, despite the provision of interventions that target the difficulties. (2) The affected academic skills are substantially and quantifiably below those expected for the individual. (3) The difficulties are not attributable to other disorders.	Specific learning disorder; Dyscalculia; Dyslexia
Motor disorder (MD)	(1) The acquisition and execution of coordinated motor skills is substantially below that expected given the individual's chronological age and opportunity for skill learning and use. (2) The deficits significantly and persistently interfere with activities of daily living. (3) The difficulties are not attributable to other disorders.	Developmental coordination disorder; Stereotypic movement disorder; Tic disorders

* Adapted from *Diagnostic and Statistic Manual of Mental Disorders, Fifth edition*.

** All subtypes of NDDs have the diagnostic criterium 'Onset during the developmental period' in addition to other criteria.

Severity scoring

For the majority of subtypes of NDDs, severity scores are available in the DSM-5 criteria.¹ For the subtypes ID and ASD, more extensive severity scoring tables have been developed.¹⁻³

For these latter disorders, severity scores are based on the level of adaptive functioning of a patient (Figure 1.1). The severity of ID is scored on functioning in three different domains: the conceptual, social and practical domains, and leads to a classification of 'Mild', 'Moderate', 'Severe' or 'Profound' ID. Severity scoring of ASD is based on the specifiers 'Social communication' and 'Restricted, repetitive behaviors'. Scoring on these domains leads to three different levels of severity: level 1: 'Requiring support', level 2: 'Requiring substantial support' and level 3: 'Requiring very substantial support'. Overall, detailed severity scoring

of NDDs is important to determine the level of support a patient requires. In DSM-5 criteria, supported by guidelines of the American Association for Intellectual and Developmental Disability, adaptive functioning and not IQ levels of patients are used for severity scoring in patients with ID.^{1,3} This is in contrast to the ICD-10 guidelines, where IQ scores determine the level of severity in a patient with ID. Using this guideline, patients with IQ levels between 50 and 69 have mild ID (developmental age from 9 to under 12 years), between 35 and 49 have moderate ID (developmental age from 6 to under 9 years), between 20 and 34 have severe ID (developmental age from 3 to under 6 years), and under 20 have profound ID (developmental age under 3 years).²

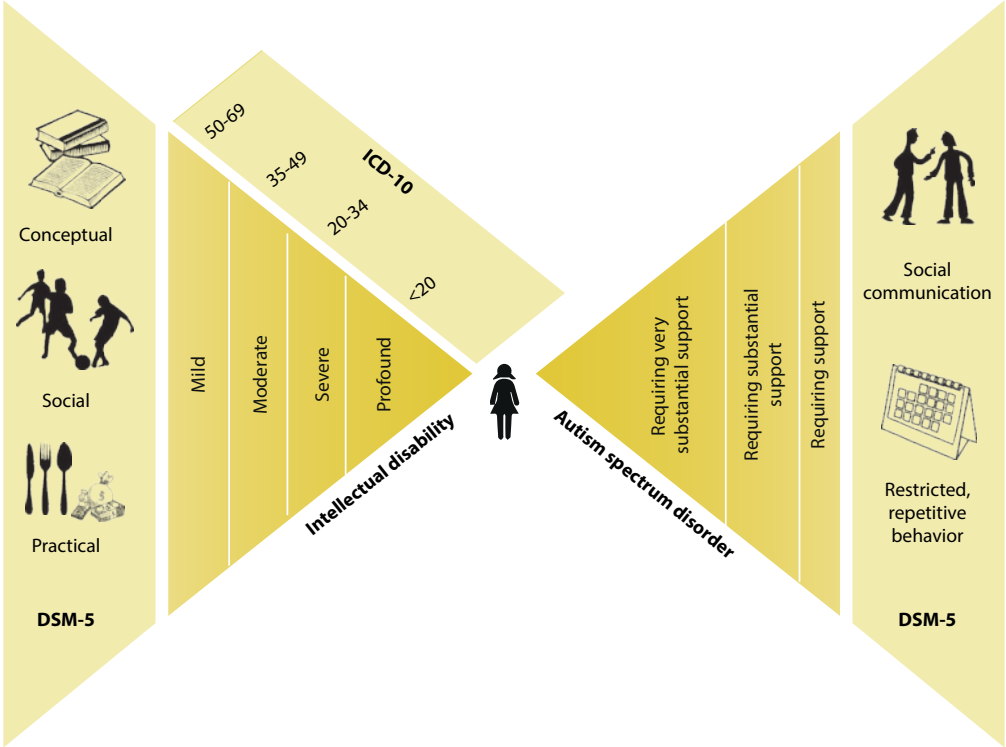


Figure 1.1: Severity scoring for intellectual disability and autism spectrum disorder based on DMS-5 and ICD-10 criteria. DSM-5 scoring elements are shown on the left for intellectual disability and right for autism spectrum disorder.

Prevalence

The overall general population prevalence of different subtypes of NDDs ranges between 1% (ID; ASD) and 10% (SLD) (Figure 1.2).⁴⁻¹² For subtypes communication disorders and motor disorders, no world-wide prevalence figure has been reported. Prevalence estimates of diagnoses within these subtypes are available, such as for language disorders (2.6-6.8%), developmental coordination disorders (5.0-5.5%), and tic disorder (0.3-1.0%).¹³⁻¹⁷ Overall, these numbers reflect the high prevalence of NDDs in the population and their large impact on health care.

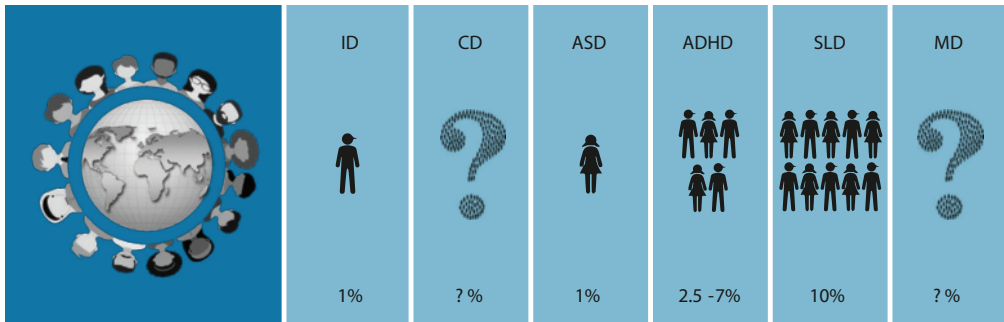


Figure 1.2: World-wide prevalence of NDD subtypes. ID = intellectual disability; CD = communication disorder; ASD = autism spectrum disorder; ADHD = attention-deficit hyperactivity disorder; SLD = specific learning disorder; MD = motor disorder

1.2 Establishing a genetic diagnosis in patients with neurodevelopmental disorders

Genetic abnormalities contribute substantially to the etiology of NDDs, with or without somatic characteristics such as epilepsy, brain abnormalities, congenital malformations and syndromic facial appearance.¹⁸⁻²¹ Therefore, genetic testing is routinely considered in the care of patients with NDDs. Establishing a genetic diagnosis is important, because identification of the genetic cause of a NDD provides parents with insight in the prognosis and disease management for their child. A genetic diagnosis may clarify the recurrence risk for parents and other family members, which facilitates decisions on future pregnancies and prenatal diagnostic testing.²² The emotional impact of a genetic diagnosis should not be underestimated. Parents often searched for a long time in order to find an explanation of the NDD in their child, a disorder that has a major impact on quality of life of not only the patient, but also parents and family.²³⁻²⁵ A genetic diagnosis allows the parents to name the disorder, share emotions and experiences with parents of other patients, become more accepting towards the situation, and to cope better with feelings of guilt.²⁶

For years, genetic testing started with a clinical diagnosis ('phenotype-first approach'): the phenotype of the patient was leading in the decision which tests should be performed. From the 1970s, chromosomal karyotyping became a first-tier test in the evaluation of all patients with NDDs. For single gene disorders, a clinical diagnosis was followed by a single, targeted test, such as testing of the *CHD7* gene on Single Nucleotide Variants (SNVs) using Sanger sequencing in patients with CHARGE syndrome, or methylation analysis of the *FMR1* repeat region for males clinically diagnosed with Fragile-X syndrome. With the introduction of the chromosomal microarray as diagnostic test, the balance started shifting away from the phenotype driven approach, towards a more genotype driven ('genotype-first') approach:

copy number variations (CNVs) such as deletions and duplications, could be found genome-wide, and enabled physicians to diagnose patients with a less distinct phenotype or suspicion for a specific syndrome. A dramatic increase in diagnostic possibilities was brought about with the introduction of Next Generation Sequencing (NGS) technologies, mainly Whole Exome Sequencing (WES) and Whole Genome Sequencing (WGS). While SNVs previously only could be identified by targeted testing of a single gene, NGS enabled the analysis of coding regions of all genes (the exome) with a single test.

In 2010, WES was effectively used in a small cohort of unselected patients with ID.²⁷ In the following years, many large-scale WES studies were performed in either research or diagnostic settings, and were focused mainly on two subtypes of NDDs: ID and ASD.²⁸⁻³³ For other subtypes of NDDs, a small number of studies has been published, such as for tic disorders (MD), ADHD, and language disorders (CD),³⁴⁻³⁸ with relatively small sample sizes (range 11-325 patients). For two other disorders, epileptic encephalopathy (EE) and schizophrenia, much larger WES studies were performed, that showed a surprisingly large overlap of genes involved in EE, schizophrenia and NDDs.³⁹⁻⁴⁵ WGS has been reported in research settings to be even more powerful than WES for the identification of genetic disorders in patients with NDDs.⁴⁶⁻⁴⁸ Diagnostic WGS is not currently available in a diagnostic setting with the exception of the United Kingdom where the 100.000 Genomes project is championing this approach.

It can be discussed that for patients with a clear clinical presentation of a disorder with low genetic and clinical heterogeneity, such as Fragile X syndrome (*FRM1* repeat expansion) or Neurofibromatosis (NF1 mutation), a targeted genetic test should precede WES as first-tier test.⁴⁹⁻⁵² But for the majority of patients the phenotype is less distinct, and for these, WES is likely to be faster and more cost-effective. Indeed, a higher diagnostic yield and lower costs than alternative diagnostic routes, including targeted gene tests, metabolic screening and brain imaging, have been reported in several studies that used WES.^{49; 53-57} Currently, early use of WES in the diagnostic trajectory of patients with NDDs is an attractive option.

1.3 Whole Exome Sequencing: a complex diagnostic work-flow

Soon after the publication of the successful results of WES for the identification of disease causing variants of patients with unexplained NDDs, this approach was introduced in diagnostic settings. The complex diagnostic WES work-flow comprises different steps, beginning and ending with the counseling of patients and their parents* (Figure 1.3).

* In this thesis, counseling of 'parents' refers to parents of patients, but also to caregivers or legal guardians of patient.

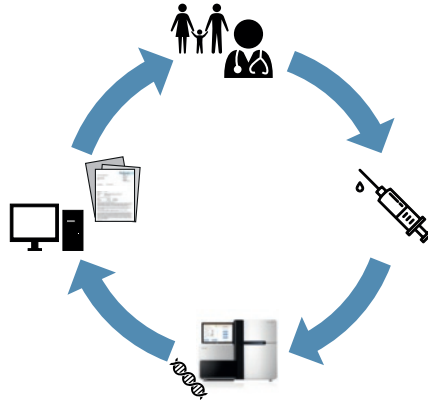


Figure 1.3: Diagnostic WES work-flow, starting with counseling of patients and parents, followed by collection of material of patients and their parents, performance of WES with variant calling and mapping, and data analysis, interpretation and classification of variants. The procedure ends where it started: the counseling of patients and their parents.

Step 1: Counseling of the patient

Patients with NDDs and their parents are informed about the opportunities, benefits and risks of genetic testing. There are usually two options in the WES diagnostic route. The physician has the opportunity to order a targeted gene panel analysis, which includes previously reported known NDD genes. Depending on the wish of parents and the center where WES is performed, analysis of the rest of the exome is offered when targeted gene panel screening reveals no diagnosis. For the latter analysis, counseling for the risk of an incidental finding is important, referring to a situation where a genetic variant is identified for a condition unrelated to the indication for testing but which might have medical value for patient care.⁵⁸ These incidental findings are identified in 2% - 4% of the WES analysis, if these variants are actively searched for.⁵⁸⁻⁶¹

Step 2: Collection of patient material

A venous blood sample is taken from patients and their parents. Some patients have clinical signs suggesting that the mutation is not present in the germline, but in a specific tissue only. Clinical symptoms such as Blaschko lines (see chapter 9) can be indicative for a mosaicism. Such somatic mosaicism can be explained by the presence of two or more populations of cells with different genotypes.⁶² Collection of additional tissues, such as skin fibroblasts, urine or buccal cells, may provide valuable diagnostic information in such cases.

Step 3: Whole exome sequencing, variant calling and mapping

The WES analysis, variant calling and mapping, are complex technological processes, that can differ between centers, depending on techniques and software used.⁶³⁻⁶⁵ Sufficient sequencing coverage of the exome and correct mapping to the reference human sequence are important aspects to obtain results of good quality.^{64; 66} To allow better mapping, the

reference genome is continuously improved by the Genome Reference Consortium. The latest version (GRCh37 or Hg38) was released in 2013 (<https://www.ncbi.nlm.nih.gov/grc>). Quality scores can be used to determine if variants identified with WES need to be validated by Sanger sequencing. Variants with good quality scores (≥ 500) reliably indicate a true variant, for which further confirmation by Sanger sequencing is not needed.^{67, 68}

Step 4: Data analysis, interpretation and classification of results

Computational (*in silico*) analysis of identified variants occurs in two steps. In the first step, both SNVs and CNVs of genes in the NDD gene panel are analyzed. If parents gave permission and, preferably, if parental DNA is available, the rest of the exome is analyzed in the second step. To detect the relevant disease-causing variants in the patient, variants are filtered on different characteristics: mode of inheritance, frequency in the population, protein function, variant type, and an overlapping phenotype with other patients who also have been detected with a variant in the same gene(s).

Mode of inheritance

Genetic variation between individuals is common: a typical human genome varies at 4.1 to 5.0 million positions compared to the human reference genome.⁶⁹ The average human genome only contains between 44 and 82 *de novo* SNVs: variants that occurred during the formation of the gametes which are therefore not present in parental DNA. Only a small part of these *de novo* variants, one to two, affect the coding sequence.⁷⁰ *De novo* mutations have been shown to represent an important cause of NDDs.²⁸⁻³² Filtering on these *de novo* mutations, as well as on maternally inherited X-linked variants in males and autosomal (bi-allelic) recessive variants, is important in the selection process of potential pathogenic variants.

Frequency in the population

From the millions of genome-wide variants present in every individual, only a small proportion are rare in the population (minor allele frequency $<0.5\%$: 40,000-200,000 variants per individual).⁶⁹ Variants that cause severe disorders such as NDDs are unlikely to become frequent in the population, given the fact that the reproductive disadvantage, or even reproductive lethality, of patients with NDDs will lead to strong negative selection. Inhouse databases and online population databases such as GnomAD (including ExAC⁷¹), contain variants identified in $>120,000$ exome sequences and $>15,000$ whole-genome sequences. These are helpful in distinguishing between common and rare variants. For interpretation of exome data, a minor allele frequency of $<1\%$ is often used as the initial cut-off value.⁷²

Protein function

For the interpretation of the estimated effect of a variant on cellular and tissue levels, it is important to evaluate the function of the protein that could be disturbed by the identified

variant.^{72; 73} Large studies on expression patterns of proteins provide insight whether a protein is involved in important cellular mechanisms in tissues such as the central nervous system, a tissue relevant for NDDs.^{74; 75} Additional information on protein function comes from *in vitro* and *in vivo* model systems such as the fruitfly, zebrafish, mouse and yeast. If a model system with disruption of a homologue gene shows a consistent phenotype with the patient, this can provide supporting evidence on the pathogenicity of the variant identified in the patient. Absence of a phenotype on disruption in a model system, however, does not exclude the gene from consideration in a given patient. Information on animal models can be retrieved from the literature, or from specialized databases such as The Mouse Genome Informatics database and the SysID database, for mice and *Drosophila* respectively.^{76; 77} Finally, variants in genes encoding a protein that has been linked to other proteins or protein pathways already implicated in NDDs, increase the likelihood that the variant is relevant in the context of explaining the patients clinical presentation.⁷⁸

Variant type

Different types of variants can be identified. SNVs change a single nucleotide, and are further classified by their impact on function as nonsense-, frameshift-, missense, and splice-site variants. CNVs are larger structural variations affecting the number of copies of a considerable number of base pairs, such as microdeletions or duplications. Nonsense-frameshift- and indel variants as well as larger structural deletions generally have a disruptive loss-of-function effect. If the loss of function of a single allele suffices to cause a phenotype in a patient, this is called haploinsufficiency. Missense variants are often more difficult to interpret, as their impact on protein function can be either through a loss-of-function (hypomorph), dominant-negative, change-of-function (neomorph) or gain-of-function (hypermorph) effect. In general, missense variants that are highly conserved among different species are more likely to be pathogenic than those that are not conserved, especially if they localize to a protein domain with known function. Further supporting evidence for the functional significance of missense variants can be provided by *in silico* prediction programs such as SIFT, Polyphen, MutationTaster and the Combined Annotation-Dependent Depletion (CADD) score.⁷⁹⁻⁸² The accuracy of these prediction programs varies between algorithms and they are often less informative than direct experimental evidence.^{73; 83} Variants in the flanking regions of the exons are known to have the potential to alter splicing. To predict whether these variants indeed lead to disrupted splicing, the Human Splicing Finder tool can be used.⁸⁴ If the gene is expressed in the tissue to be tested, investigation of patient mRNA can be valuable to verify pathogenicity of these variants.

Overlapping phenotype with other patients with (possible) pathogenic variants in the same gene

For interpretation of variants, comparison of clinical features observed in the patient with the phenotype associated with the disorder is essential. If the clinical features in the patient

strongly overlap with those of other patients, the variant is more likely to be pathogenic. Conversely, if the phenotype of the patient differs from the clinical spectrum reported in the literature, it is less likely that the identified variant is clinically relevant. This criterion is not absolute since it is possible that a novel genotype-phenotype correlation is present.

After extensive analysis of WES data, identified variants are classified according to appropriate guidelines.^{85; 86} A diagnostic report with the results and conclusion is written by the clinical laboratory specialists and sent to the physician who ordered the test.

Step 5: Counseling of the patient

Patients and their parents are informed by their physician about the diagnostic findings:

1. A conclusive diagnosis is found: a pathogenic (class 5) or likely pathogenic (class 4) variant is identified in a gene previously associated with the phenotype of the patient.
2. A possible diagnosis is reported: a variant of unknown significance (class 3) is present in a gene associated with the patient's phenotype, or a pathogenic variant (class 4 or 5) is identified in a gene not related to a disorder before. Such variants can be shared in matchmaker programs such as Genematcher, Matchmaker exchange (<http://www.matchmakerexchange.org/>) or PhenomeCentral to find additional patients, and can finally lead to the identification of novel NDD genes.^{87; 88}
3. No diagnosis is identified: no variants that may explain the disorder are observed, or only likely benign (class 2) or benign (class 1) variants in known genes are identified.

Prognostic and therapeutic consequences of the diagnosis should be discussed. This is often more difficult for possible diagnoses than for conclusive diagnoses: the novelty of the diagnosis is reflected in a lack of knowledge among caregivers. Patients and parents experience difficulty in finding information and peer support.²⁶ In addition to the prognostic and therapeutic consequences, counseling of the recurrence risk is important. If the variant occurred *de novo*, there is a low recurrence risk for parents to have a second child with the same disorder due to germline mosaicism in one of the parents (~0.1-1.3%).^{89; 90} This risk is higher if one of the parents carries the variant in mosaic form^{90; 91} or for maternal carriers of an X-linked variant (50% recurrence risk for male offspring; 50% carrier risk for female offspring). Recurrence risk equals the population risk for the estimated 6.5-7.5% of patients in whom the mutation occurred post-zygotically.^{92; 93}

1.4 Diagnostic yield of WES

Studies applying WES to cohorts of patients with NDDs, report an overall diagnostic yield of 16%-32% (conclusive diagnosis) and an increase of 4%-24% by adding patients with a possible diagnosis.^{28; 29; 53; 94-100} The diagnostic yield for a conclusive and possible diagnosis together ranged between 25% (published in 2013) and 56.6% (published in 2017)^{53; 96} The largest study investigating >4000 individuals with NDDs (2017), reports an estimated 42% of patients to carry a pathogenic mutation.³³ The increased diagnostic yield over time can be explained by improvement in WES techniques and interpretation of variants due to the increased use of matchmaker programs and the growing list of genes linked to NDDs.¹⁰¹ In addition to SNV analysis, exome-based read-depth CNV screening increases the diagnostic yield with ~2%.¹⁰² Re-evaluation of WES-data by inspection of recently published literature, the use of new bioinformatic pipelines, and critical assessment of the phenotype by a medical geneticist, can further increase the diagnostic field of WES further with 4.2% to 15.4%.¹⁰³⁻¹⁰⁸ WES is not only effective in children, but also in adults with various forms of NDDs: in 27.7% of undiagnosed adults, a diagnosis was found.¹⁰⁹ High diagnostic yields are also observed in cohorts of abnormal fetuses. Studies report a diagnosis in 24% of deceased fetuses and a conclusive or possible diagnosis in 47% of prenatal WES on fetuses with developmental anomalies.¹¹⁰⁻¹¹² Targeted testing using a WES gene panel in probands without the availability of DNA from parents (singleton testing) can still be diagnostic although it has a lower diagnostic yield (11%).¹¹³



2

Scope and outline of this thesis

2.1 The scope of this thesis

With a diagnostic yield up to 50% for patients with unexplained NDDs, WES has proven successful as a diagnostic genetic test. It led to the identification of hundreds of novel NDD genes, many of which are associated with unique neurodevelopmental syndromes. Still, a significant proportion of NDD patients with rare genetic disorders remains undiagnosed. For those patients molecularly diagnosed with a NDD that was recently discovered by WES, available information on clinical, prognostic and management characteristics is often limited. To provide further insight into the genetic, clinical, and biological aspects of NDDs in these patients, I aimed in this thesis to (Figure 2.1):

- Discover rare NDDs by (1) performing WES on large scale, (2) using different approaches to interpret WES data, and (3) inclusion of mildly affected patients for WES, a less frequently investigated subset of NDD patients.
- Confirm and delineate associated phenotypes of newly identified disorders by (1) using matchmaker systems and (inter)national collaborations, and (2) collecting and comparing detailed phenotypic data.
- Confirm newly identified disorders and investigate underlying biological mechanisms by using functional experiments.

2.2 The outline of this thesis

We established a cohort of 826 patients (56% male; 44% female; median age 8 years with interquartile range of 4-16 years) with unexplained ID, defined as the presence of (1) deficits in intellectual functioning; (2) deficits in adaptive functioning, and (3) onset during developmental period.¹ This cohort was composed as a follow-up cohort to previously clinically well-defined cohorts (250k SNP array: 1,388 patients¹⁹; WES: 10 patients²⁷; WES: 100 patients²⁸; WGS: 50 patients⁴⁶). WES was performed in all patients and their parents and results were compared with WES results of patients with NDDs who were sequenced in other (inter)national centers. For 783 patients (95%) of our cohort, additional detailed clinical data were available.

Cohort (Figure 2.2)

Neurodevelopmental profile

ID was present in all patients and severity varied: 28.0% had mild ID, 20.7% had moderate ID, 28.2% had severe ID. For the remaining 23.1%, severity scoring was not performed or not possible due to the young age of the patient. In 32.2% of the patients, behavioral problems such as NDD subtypes ASD and ADHD were reported. Information on diagnoses of the NDD subtypes CM, SLD, or MD was generally not available.

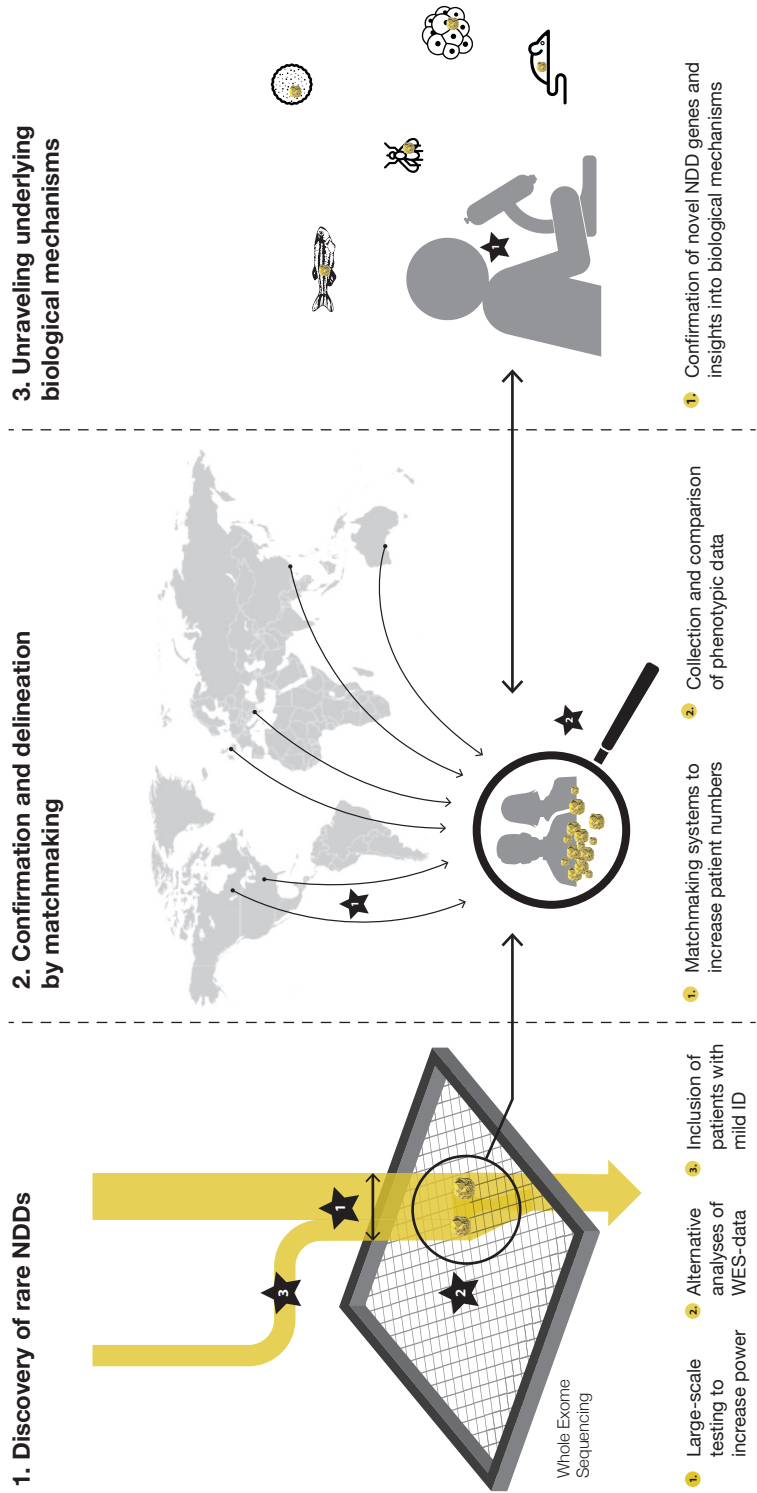


Figure 2.1: Illustration of the scope of this thesis. (1) To discover rare NDDs, a large cohort of patients (yellow arrow) is tested by WES (the sieve), and mutations causing novel NDDs (gold nuggets) can be filtered out using different approaches to interpret WES data. Next, I aim to (2) confirm and delineate these novel NDDs by collection of additional patients using matchmaking systems, and (3) by testing of mutations in novel NDD genes in model systems.

Clinical features

The majority of patients (739/783, 94%) had at least one additional feature. The most frequent co-morbid features were neurological problems: epilepsy (24.5%), hypotonia (20.3%), brain abnormalities (19.8%), spasticity (7.0%), cerebral visual impairment (4.7%) or hypertonia (3.2%). Other frequently observed features included gastrointestinal abnormalities (23.1%), eye abnormalities (22.3%), skeletal abnormalities (14.7%), urogenital malformations (9.9%), cardiac malformations (7.4%), and cleft lip and/or palate (2.1%). Growth was abnormal in 38.6% of the patients: 16.9% had abnormal height (13.9% short stature and 3.0% tall stature), 14.1% had abnormal weight (6.1% low weight and 8.0% obesity), and 20.7% had aberrant head circumference (6.4% macrocephaly and 14.3% microcephaly). The observed frequencies of these features were comparable to previously described cohorts of patients with ID or NDDs.^{18; 19; 28}

Results

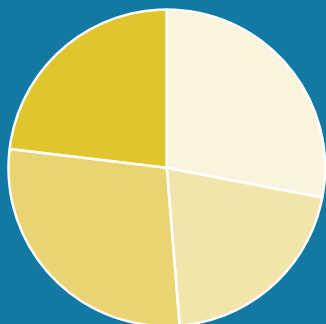
Interesting findings of WES in this cohort, combined with results of matchmaking, clinical evaluation and functional experiments, are reported in **Chapter 3 – 11**.

In **Chapter 3**, we report the identification of candidate ID and NDD genes, by statistical meta-analysis of *de novo* mutations from exome data of patients in our cohort and *de novo* mutations in patients with ID (2,104 patients) or other NDD subtypes (4,102 patients) published in the literature. These analyses resulted in the discovery of ten candidate ID genes and showed us that there is considerable overlap between patients with ID and patients with other subtypes of NDDs.

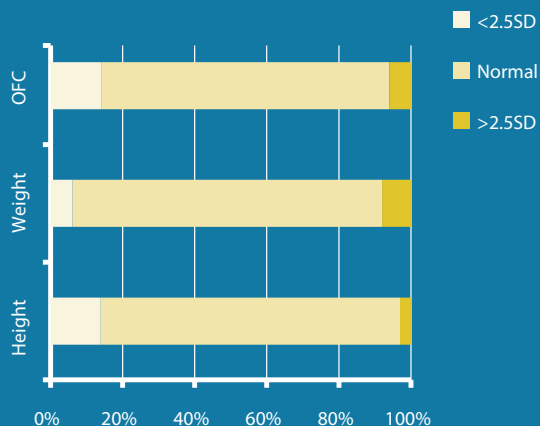
In the next three chapters, we further investigated three novel candidate ID and NDD genes: *SON*, *TLK2* and *RAC1*. We compared detailed phenotype data of patients with *de novo* mutations in these genes and performed different types of functional studies to investigate underlying disease mechanisms. In **Chapter 4**, we report that 20 patients with *de novo* mutations in *SON*, have a severe developmental phenotype and malformations of the brain cortex. At a cellular level, we showed that mutations in *SON* led to disturbed splicing of downstream target genes, partly explaining the phenotype observed in the patients. In **Chapter 5**, a large number of individuals (n=40), mostly selected through international matchmaking, are reported to have autosomal dominant mutations in *TLK2*. The developmental spectrum of patients with a mutation in this gene is mild and comprises a broader NDD spectrum. The large number of patients enabled us to define this clinical spectrum in detail. **Chapter 6** describes specific *de novo* missense mutations in the small GTPase *RAC1* in seven patients. Patients with this novel syndrome show remarkable differences in head size. Using mouse fibroblasts and zebrafish as model systems, we

Level of intellectual disability

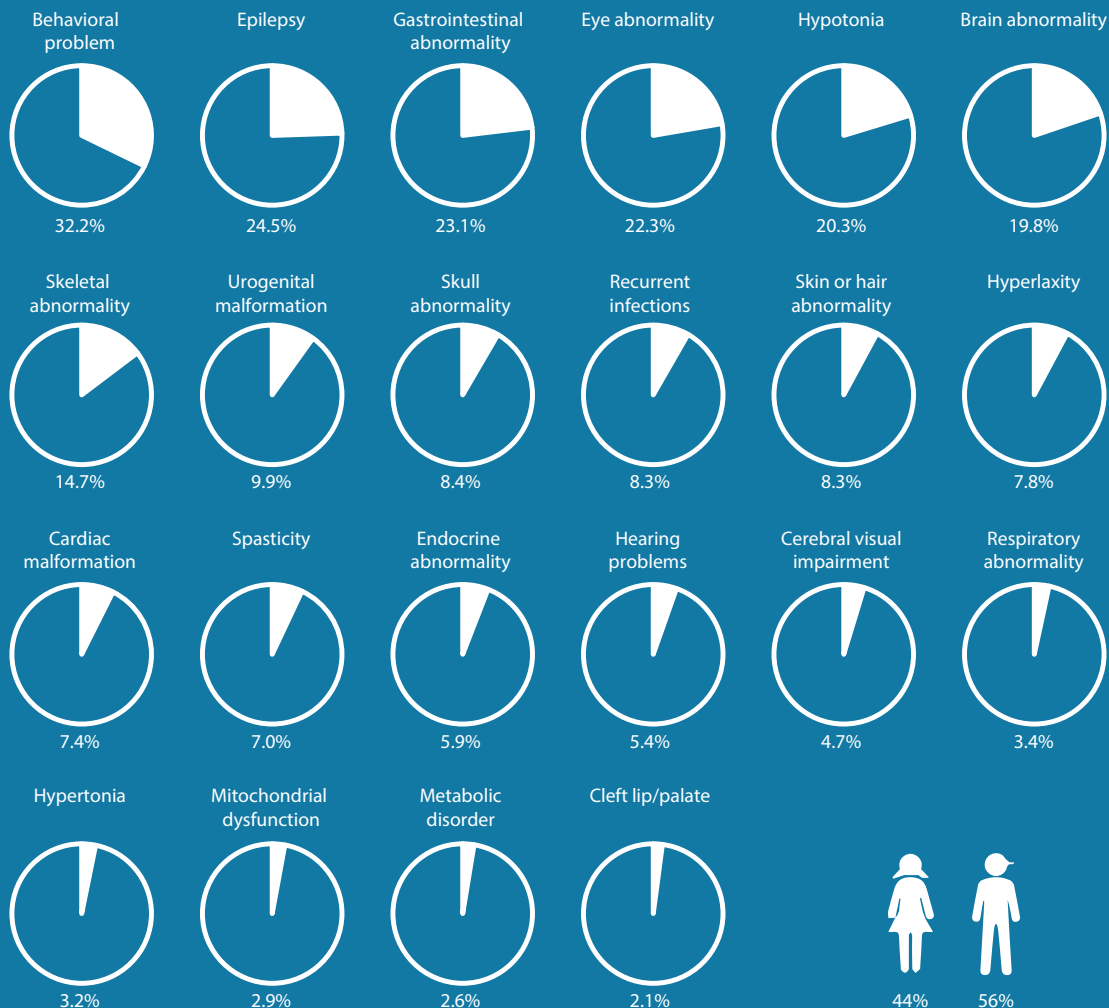
- Mild
- Moderate
- Severe
- Unknown



Growth parameters



Other features



provide more insight in the underlying mechanisms of the observed phenotypic differences between patients.

In **Chapter 7**, *de novo* mutations in WES data of our patient cohort are analyzed using a pathway approach. Genes related to the mTOR pathway were selected and evaluated for the presence of *de novo* mutations. We show that the majority of patients with mutations in this pathway have a large head circumference: macrocephaly occurs significantly more frequently than in patients without an mTOR-related mutation. In line with this observation, we report that the same set of mTOR-related genes significantly contributes to intracranial volume in the general population. Finally, we describe three patients with severe ID, epilepsy and macrocephaly with mutations in the mTOR activator *RHEB*, and show in zebrafish and mouse models that these clinical features reflect altered mTOR pathway activity. In **Chapter 8**, we report two specific missense mutations in the *RAB11B* small GTPase in five unrelated patients and a similar brain phenotype. Functional studies further support the pathogenicity of these two recurrent mutations. **Chapter 9** describes the discovery of a distinct NDD syndrome in females, caused by *de novo* loss-of-function mutations in the X-linked gene *USP9X*. Females with this syndrome have a variety of congenital malformations and brain malformations. We show the recognizability of the syndrome by expanding the initial cohort of 13 patients, established by extensive international collaborations, with four patients identified through targeted testing. **Chapter 10** reports on 10 *de novo* loss-of-function mutations in *WAC*, collected through searches in matchmaking databases and (inter)national collaborations. *WAC* haploinsufficiency causes a NDD syndrome and partially explains the phenotype observed in patients with 10p12p11 microdeletions. We also show that loss-of-function of *WAC* in *Drosophila* leads to learning difficulties, similar to patients. In **Chapter 11**, we delineate the clinical spectrum of the recently discovered PURA syndrome, caused by *de novo* mutations in *PURA*. In total, we collected clinical data for 32 patients from 24 different diagnostic centers in nine different countries. A genotype-phenotype analysis revealed that the disease severity is not related to the type and position of the mutation.



3

Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability

This chapter has been published as:

Stefan H. Lelieveld*, **Margot R.F. Reijnders***, Rolph Pfundt, Helger G. Yntema, Erik-Jan Kamsteeg, Petra de Vries, Bert. B.A. de Vries, Marjolein H. Willemsen, Tjitske Kleefstra, Katharina Löhner, Maaïke Vreeburg, Servi Stevens, Ineke van der Burgt, Ernie M.H.F. Bongers, Alexander P.A. Stegmann, Patrick Rump, Tuula Rinne, Marcel R. Nelen, Joris A. Veltman, Lisenka E.L.M. Vissers*, Han G. Brunner*, Christian Gilissen*

Nature Neuroscience (2016) 19, 1194-1196

* These authors contributed equally

Abstract

To identify candidate genes for intellectual disability, we performed a meta-analysis on 2,637 *de novo* mutations, identified from the exomes of 2,104 patient–parent trios. Statistical analyses identified 10 new candidate ID genes: *DLG4*, *PPM1D*, *RAC1*, *SMAD6*, *SON*, *SOX5*, *SYNCRIP*, *TCF20*, *TLK2* and *TRIP12*. In addition, we show that these genes are intolerant to nonsynonymous variation and that mutations in these genes are associated with specific clinical ID phenotypes.

Report

Intellectual disability (ID) and other neurodevelopmental disorders are in part due to *de novo* mutations affecting protein-coding genes.^{31; 40; 46; 114} Large scale exome sequencing studies of patient–parent trios have efficiently identified genes enriched for *de novo* mutations in cohorts of individuals with ID compared to controls³¹ or on the basis of expected gene-specific mutation rates.³²

Here we sequenced the exomes of 820 patients with ID and their parents as part of routine genetic testing at the Radboud University Medical Center (RUMC) in the Netherlands. We identified 1,083 *de novo* mutations (DNMs) in the coding and canonical splice site regions affecting 915 genes (Supplemental Tables 3.1 and 3.2 and Supplemental Figures 3.1–3.4). In our cohort we detected an increased number of loss-of-function (LoF) mutations compared to controls (Fisher's exact test, $P = 9.38 \times 10^{-12}$; Methods) and enrichment for recurrent gene mutations (observed versus expected, $P < 1 \times 10^{-5}$; Supplemental Figure 3.5).

Using an established framework of gene specific mutation rates,¹¹⁵ we calculated for each gene the probability of identifying the observed number of LoF or functional DNMs in our cohort (Methods). To validate this approach we first performed the analysis on the complete set of 820 ID patients. After Benjamini–Hochberg correction for multiple testing, 18 well-known ID-associated genes were significantly enriched for DNMs (Supplemental Tables 3.3 and 3.4). To optimize our analysis for the identification of new candidate genes in the RUMC cohort, we excluded all individuals with mutations in any of the known ID genes (Material and methods and Supplemental Figure 3.6). Repeating the analysis for mutation enrichment, we identified four genes (*DLG4*, *PPM1D*, *SOX5* and *TCF20*) that were not, to our knowledge, previously associated with ID and that were significantly enriched for DNMs in our cohort (Figure 3.1, Table 3.1 and Supplemental Table 3.5). To achieve the best possible power for the identification of candidate ID genes, we next added data from four previously published family-based sequencing studies (Supplemental Table 3.1). The combined cohort included 2,104 patient–parent trios and 2,637 DNMs across 1,990 genes. After again excluding individuals with mutations in known ID genes, this cohort consisted of 1,471 individuals with 1,400 DNMs in 1,235 genes (Material and methods and Supplemental Figure 3.6). Meta-analysis on this combined cohort identified ten candidate ID genes with more LoF DNMs or more functional DNMs than expected *a priori*. These ten genes included the four candidate ID genes previously identified in the RUMC cohort, as well as *RAC1*, *SMAD6*, *SON*, *TLK2*, *TRIP12* and *SYNCRIP* (Figure 3.1, Table 3.1 and Supplemental Table 3.6).

To further evaluate the identification of the ten candidate ID genes, we compared the phenotypes of the 18 RUMC individuals with DNMs in these genes. We observed strong phenotypic overlap for some of these genes (Figure 3.2, Supplemental Table 3.7 and Supplemental Note). Additional genes that were close to statistical significance, such as

Table 3.1 Candidate ID genes

Gene	RUMC cohort		ID cohort		Gene description
	LoF	Functional	LoF	Functional	
DLG4 NM_001365.3	q 1.13E-04	0.086	7.69E-06	8.02E-04	Required for synaptic plasticity associated with NMDA receptor signaling. Depletion of DLG4 changes the ratio of excitatory to inhibitory synapses in hippocampal neurons.
	p 6.56E-09	7.50E-06	2.24E-10	4.66E-08	
	c (n=3)	(n=3)	(n=4)	(n=5)	
PPM1D NM_003620.3	q 0.047	0.764	8.22E-04	0.174	Ser/Thr phosphatase which mediates a feedback regulation of p38-p53 signaling thereby contributing to growth inhibition and suppression of stress induced apoptosis.
	p 5.45E-06	2.56E-04	9.57E-08	3.03E-05	
	c (n=2)	(n=2)	(n=3)	(n=3)	
RAC1 NM_018890.3	q	0.217		0.020	Plasma membrane-associated small GTPase involved in many cellular processes. In the synapses, it mediates the regulation of F-actin cluster formation by SHANK3.
	p n.d.	3.80E-05	n.d.	1.75E-06	
	c	(n=2)		(n=3)	
SMAD6 NM_005585.4	q		0.037	1	Mediates TGF-beta activity and the BMP-SMAD1 signaling. Functions as a transcriptional co-repressor.
	p n.d.	n.d.	8.29E-06	7.50E-04	
	c		(n=2)	(n=2)	
SON NM_138927.1	q 1	1	0.003	1	Component of the spliceosome that plays pleiotropic roles during mitotic progression. Functions in efficient cotranscriptional RNA processing.
	p 0.086	0.005	4.12E-07	1.67E-03	
	c (n=1)	(n=1)	(n=3)	(n=3)	
SOX5 NM_006940.4	q 0.016	1	0.038	0.216	Member of Transcription factors that regulate embryonic development. Plays a critical role in neuronal progenitor development by regulating the timing of differentiation.
	p 1.39E-06	3.98E-04	8.79E-06	5.83E-05	
	c (n=2)	(n=2)	(n=2)	(n=3)	
SYNCRIP NM_006372.4	q 1	1	0.028	1	Heterogeneous nuclear ribonucleoprotein (hnRNP) functioning in the CRD-mediated mRNA stabilization complex and the SMN complex, and the apoB RNA editing-complex.
	p 0.001	0.019	4.94E-06	1.24E-03	
	c (n=1)	(n=1)	(n=2)	(n=2)	
TCF20 NM_005650.1	q 6.22E-06	0.035	1.24E-04	0.174	Transcriptional activator of matrix metalloproteinase 3 and (co)activator of various other transcriptional activators.
	p 1.81E-10	1.00E-06	7.21E-09	3.71E-05	
	c (n=4)	(n=4)	(n=4)	(n=4)	
TLK2 NM_005650.1	q 0.100	1	0.347	5.86E-04	Ser/Thr kinase regulating chromatin assembly. Involved in DNA replication, transcription, repair and chromosome segregation.
	p 1.44E-05	4.20E-04	9.09E-05	1.70E-08	
	c (n=2)	(n=2)	(n=2)	(n=5)	
TRIP12 NM_001284214.1	q 0.273	1	2.35E-04	0.174	E3 ubiquitin-protein ligase involved in ubiquitin fusion degradation pathway. Guards excessive spreading of ubiquitinated chromatin at damaged chromosomes in DNA repair.
	p 5.55E-05	0.003	2.05E-08	4.05E-05	
	c (n=2)	(n=2)	(n=4)	(n=4)	

All genes listed reached statistical significance after Benjamini–Hochberg correction for enrichment of functional and/or LoF DNM in the RUMC or ID cohort. For each gene the Benjamini–Hochberg corrected P-value (q), uncorrected P-value (p) and raw counts (c) are shown. Significant Benjamini–Hochberg-corrected P-values are depicted in bold. n.d. (not defined) indicates genes without observed DNMs in the RUMC or ID cohort.

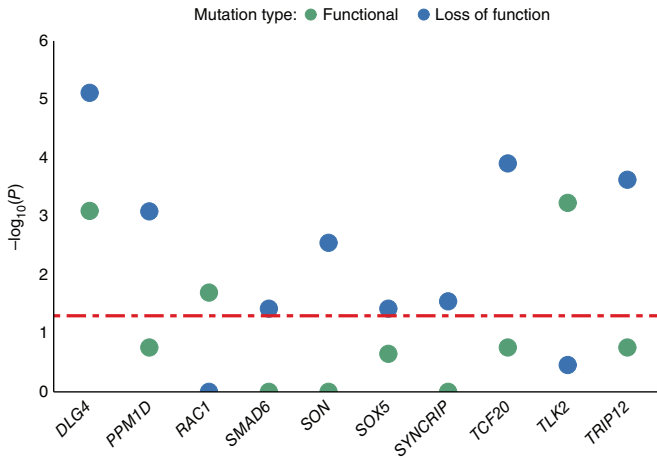


Figure 3.1: Genes enriched for LoF and functional DNMs in a cohort of 2,104 ID trios from multiple studies. The y axis shows the $-\log_{10}$ -transformed, corrected P-value of the DNM enrichment as listed in Table 1. Corrected P-values based on LoF mutations are blue and corrected P-values based on functional mutations are green. Only genes with corrected P-values (LoF, functional, or both) less than the significance threshold (red dotted line, 0.05) are shown.

SETD2, show phenotypic similarities suggestive of a shared genetic cause consistent with previous case reports (Supplemental Figure 3.7 and Supplemental Note).^{116; 117}

Studies have shown that genes involved in genetic disorders exhibit strongly reduced tolerance to nonsynonymous genetic variation compared to non-disease-associated genes. This is particularly evident for ID.⁴⁶ We found that a large set of well-known dominant ID genes ($n = 444$), along with the ten candidate ID genes, are highly intolerant of LoF variation⁷¹ (median probability of being LoF-intolerant (pLI) of 0.95, $P < 1 \times 10^{-5}$ and median pLI of 0.99, $P < 1 \times 10^{-5}$, respectively; Material and methods, Supplemental Figure 3.8 and Supplemental Table 3.8). We noted that those ID genes that harbor only missense variants ('missense-only' genes) are among the most intolerant ID genes (median pLI of 0.99, $P < 1 \times 10^{-5}$; Supplemental Figure 3.8). Additionally, we found that mutations in missense-only genes are more likely to cluster than mutations in genes for which we also identified LoF mutations ($P = 0.01$, Fisher's exact test; Material and methods and Supplemental Table 3.9).

There is considerable overlap of genes and molecular pathways involved in neurodevelopmental disorders (NDDs), such as autism spectrum disorder, schizophrenia, epileptic encephalopathy and ID.⁴⁵ Therefore, we performed a third analysis including 12 published family-based sequencing studies of various NDDs (Supplemental Table 3.1 and Supplemental Figure 3.6). Repeating our analysis in this NDD cohort, we identified seven genes significantly enriched for either LoF or functional DNMs (Supplemental Figure 3.9 and Supplemental Table 3.10). In line with our hypothesis, five of these identified genes were also identified in our previous analyses with individuals with ID only, whereas two genes (*SLC6A1*

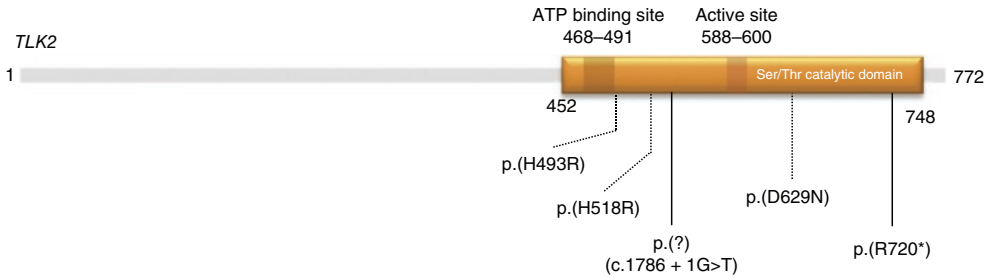


Figure 3.2: TLK2 protein (Q86UE8) with DNMs localized to the serine/threonine catalytic domain.

Two individuals in the RUMC cohort were found to have a DNM in TLK2; they showed overlapping clinical features including facial dysmorphisms (Supplemental Note).

and *TCF7L2*) only reached significance in the NDD meta-analysis as a result of additional mutations in patients with phenotypes other than ID (Supplemental Table 3.11). Specifically, for two of the five candidate ID genes (*TLK2* and *TRIP12*) additional DNMs were identified in individuals with autism spectrum disorder and schizophrenia, suggesting that DNMs in these genes may lead to a broader phenotype than ID alone. For *TRIP12*, a similarly broad phenotype has been reported previously.¹¹⁴

In summary, we identified ten candidate ID genes via a meta-analysis of whole exome sequencing data on 2,104 ID trios. The statistical framework used here differs from other methods based on gene-specific mutation rates by removing all trios with mutations in known disease genes and by applying Benjamini–Hochberg correction for multiple testing. Our study underscores the impact of DNMs on a continuum of neurodevelopmental phenotypes that impinge on a broad range of processes, including chromatin modifiers (*TRIP12* and *TLK2*), Fragile X Mental Retardation Protein (FMRP) target and synaptic plasticity genes (*DLG4*; Supplemental Figure 10) and embryonically expressed genes (*PPM1D* and *RAC1*).³¹ Data from a similar systematic study of DNMs in neurodevelopmental disorders suggest that many, and possibly most, genes whose DNM causes severe developmental disorders are now known.³³ Yet only *TCF20* and *PPM1D* are shared between the 10 candidate genes in our study and the 14 genes identified by McRae *et al.*³³ Thus, a large number of rare dominant developmental disorder genes may remain to be identified.

Material and methods

Recruitment of individuals with ID

The Department of Human Genetics from the Radboud University Medical Center (RUMC) is a tertiary referral center for clinical genetics. Approximately 350 individuals with unexplained intellectual disability (ID) are referred annually to our clinic for diagnostic evaluation. Since September 2011 whole exome sequencing (WES) has been part of the routine diagnostic work-up aimed at the identification of the genetic causes underlying disease.¹¹⁸ For individuals with unexplained ID, a family-based WES approach is used which

allows the identification of DNMs as well as variants segregating according to other types of inheritance, including recessive mutations and maternally inherited X-linked recessive mutations in males.²⁸ For this study, we selected all individuals with ID who had family-based WES using the Agilent SureSelect v4 enrichment kit combined with sequencing on the Illumina HiSeq platform in the time period 2013–2015. This selection yielded a set of 820 individuals, including 359 females and 461 males. The level of ID ranged between mild (IQ 50–70) and severe–profound (IQ <30).

Families gave informed consent both for the diagnostic procedure and for forthcoming research that could result in the identification of new genes underlying ID by meta-analysis, as presented here. Explicit consent for photo-publication was sought and granted by a subset of families.

Diagnostic whole exome sequencing

The exomes of 820 patient–parent trios were sequenced, using DNA isolated from blood, at the Beijing Genomics Institute (BGI) in Copenhagen. Exome capture was performed using Agilent SureSelect v4 and samples were sequenced on an Illumina HiSeq instrument with 101-bp paired-end reads to a median coverage of 75×. Sequence reads were aligned to the hg19 reference genome using BWA version 0.5.9-r16. Variants were subsequently called by the GATK unified genotyper (version 3.2-2) and annotated using a custom diagnostic annotation pipeline. Base-pair resolution coverage of the regions enriched by the SureSelect V4 kit were computed by BEDTools based on the regions as provided by the manufacturer. An average of 98.9% of Agilent SureSelect V4 enriched targets was covered by 10 or more reads for the RUMC cohort of 820 ID patients (Supplemental Figure 3.1).

Identification of DNMs in 820 individuals with ID

The diagnostic WES process as outlined above only reports (*de novo*) variants that can be linked to the individuals' phenotypes. In this study, we systematically collected all DNMs located in the coding sequence (RefSeq) and/or affected canonical splice sites (canonical dinucleotides GT and AG for donor and acceptor sites; Supplemental Figure 3.4), as identified in the 820 individuals with ID irrespective of their link to disease, to evaluate the potential relevance of genes for ID in an unbiased fashion using a statistical framework. DNMs were called as described previously.²⁸ Briefly, variants called within parental samples were removed from the variants called in the child. For the remaining variants, pileups were generated from the alignments of the child and both parents. Based on pileup results, variants were then classified into the following categories: 'maternal' (identified in the mother only), 'paternal' (identified in the father only), 'low coverage' (insufficient read depth in either parent), 'shared' (identified in both parents) and 'possibly *de novo*' (absent in the parents). Variants classified as possibly *de novo* were included in this study.

We applied various quality measures to ensure that only the most reliable variant calls were included in the study: (i) all samples had fewer than 25 possibly *de novo* calls; (ii) each variant had at least 10× coverage in either parent (for example, high prior probability of being inherited); (iii) the location was not in dbSNP version137 (for example, a possible highly mutable genomic location) and (iv) each variant was called in a maximum of 5 samples in our in-house variant database (which eliminated variants that occur too frequently to be disease-causing given the incidence of ID in combination with the sample size of our in-house database); (v) each variant showed a variant read percentage >30%, or alternatively, >20% with >10 individual variant reads; and (vi) each variant had a GATK quality score of >400. For *de novo* variants called within a 5-bp window of each other within the same individual, variant calls were manually curated and merged into a single call (when occurring on the same allele). This set of criteria resulted in the identification of 1,083 potential DNMs in 820 individuals with ID.

Validation and categorization of DNMs

In a separate (unpublished) in-house study, we recently determined the predictive value for GATK quality scores in terms of the variant being validated by Sanger sequencing. A set of 840 variants called by the same version of GATK was retrospectively analyzed for the quality scores and validation statuses of each variant in the set. Based on this assessment, we determined that a GATK quality score ≥ 500 resulted in 100% of variants being validated by Sanger sequencing (data not shown). In addition to our in-house study, two other studies also found 100% Sanger validation rates for variants with GATK quality scores of ≥ 500 .⁶⁷ Based on these results, we considered all variants with a GATK Q-score of ≥ 500 ($n = 1,039$) to be true DNMs. Nonetheless, a random set of 141 potential DNMs with GATK Q-scores of ≥ 500 were all confirmed by Sanger sequencing. All potential *de novo* variants with GATK Q-scores between 400 and 500 ($n = 40$) were subsequently validated by Sanger sequencing, and all were confirmed. All 20 DNMs of the reported candidate genes were confirmed by Sanger sequencing (Supplemental Table 3.2 and Supplemental Figures 3.2-3.4).

For further downstream statistical analysis (see below), DNMs were categorized by mutation type: (i) LoF DNM ($n = 211$), including nonsense ($n = 77$), frameshift ($n = 97$), canonical splice site ($n = 27$), start loss ($n = 2$), stop loss ($n = 1$) and premature stop codon resulting from an indel ($n = 7$); and (ii) functional DNM ($n = 872$), including all LoF mutations ($n = 211$), in-frame insertion/deletion events ($n = 23$) and all missense mutations ($n = 638$) (Supplemental Figure 3.4). For variants within the same individual and within the same gene but more than 5 bp apart, the variant with the most severe functional effect was considered for the per-gene statistics (see below).

*Evaluating the number of recurrently LoF and functional *de novo* mutated genes*

We simulated the expected number of recurrently mutated genes by redistributing the

observed number of mutations at random over all genes based on their specific LoF and functional mutation rates (see “Statistical enrichment of DNMs” below) as described by Samocha *et al.*¹¹⁵ Based on results from 100,000 simulations, we calculated how many times the number of recurrently mutated genes was the same as or exceeded the observed number of recurrently mutated genes in the RUMC data set. We performed these simulations separately for LoF and functional DNMs (Supplemental Figure 3.5). *P*-values were then calculated by taking the number of times the number of recurrently mutated genes exceeded the observed number of recurrently mutated genes and dividing by the number of simulations. In addition, *z*-values were computed by subtracting the mean value of the simulations from the observed value and dividing by the s.d. of the simulations.

Genes previously implicated in ID etiology

To evaluate whether the genes identified by our meta-analyses had been previously implicated in ID, we used two publicly available repositories of genes known to be involved in ID. First, we used our list of 707 genes, routinely used by our diagnostic setting to interpret WES results of individuals with ID.¹¹⁹ Second, we downloaded a list of 1,424 genes associated with developmental disorders from the DDG2P database (<http://www.ebi.ac.uk/gene2phenotype/gene2phenotype-webcode/cgi-bin/handler.cgi#>); the list was compiled and curated by clinicians as part of the Deciphering Developmental Disorders (DDD) study to facilitate clinical feedback of likely causal variants.³² In total the two lists comprised 1,537 unique genes. In this manuscript, the list of unique gene entries is referred to as “known ID genes” (Supplemental Table 3.4).

Statistical enrichment of DNMs

In our meta-analysis for ID and neurodevelopmental disorders we only included studies with minimum of 50 trios. For each gene, and each of the functional classes (LoF and functional), we used the corresponding gene-specific mutation rate (GSMR) as published by Samocha *et al.*¹¹⁵ to calculate the probability of the number of identified DNMs in our cohort. For genes for which no GSMR was reported, we used the maximum GSMR of all reported genes (i.e., the GSMR of the gene *TTN*). We then calculated specific mutation rates for the two defined functional classes (LoF, functional). The GSMR for LoF DNMs was calculated by summing the individual GSMR for nonsense, splice site and frameshift variants; the GSMR for functional DNMs was calculated by summing the GSMR for the LoF variants with the missense mutation rate; and for genes for which variants from different functional classes were identified, we used the overall GSMR. For the stop-loss and start-loss mutations we used the LoF-rate and for in-frame indels, the functional rate. Null hypothesis testing was done using a one-sided exact Poisson test based on a sample size of 820 individuals with ID, representing 1,640 alleles for autosomal genes and 1,179 alleles for genes on the X chromosome (461 males).

For DNMs on chromosome X the correct mutation rate depends on the patient's gender as the mutation rate for fathers is higher than for mothers. Estimates show a 4:1 ratio of paternal to maternal DNMs.¹²⁰ Hence, male offspring, receiving their chromosome X exclusively from the mother, have therefore a lower mutation rate on chromosome X than estimated by the GSMR. This correction could, however, only be performed for the RUMC cohort, as gender information was not available for all studies included in the ID cohort. Notably, not correcting for this bias in male individuals for DNM in genes on the X chromosome will lead to less significant *P*-values for genes on the X chromosome, thereby potentially underestimating the significance of candidate ID genes located on the X chromosome. When a patient was found to have two DNMs in the same gene we ignored one of the two DNMs for the statistical enrichment analysis to avoid false positive results. In such cases the severity of the DNM protein effect was a factor in the choice of which DNM to ignore. For example, if a patient had one missense and one nonsense DNM in the same gene, the missense mutation was ignored in the statistical analysis.

Gene specific *P*-values were corrected for multiple testing based on the 18,730 genes present in the Agilent V4 exome enrichment kit times the number of tests (2), using the Benjamini–Hochberg procedure with an FDR of 0.05. In our cohort of 820 individuals with ID, conclusive diagnoses were already made based on DNMs in genes previously implicated in disease. The use of a multiple testing correction with a FDR of 0.05, in combination with a potential large number of DNMs in known ID genes, may artificially increase the significance of other genes because of an increasingly lenient correction for the least significant genes.¹²¹ To verify that the identification of candidate ID genes was not inflated by this effect, we performed the analysis after removing all individuals with a DNM in one of the known genes (potential other DNMs in such individuals were also removed for further analysis). Incidentally, this also increased our statistical power. The mode of inheritance was not taken into account when removing individuals with a DNM in a known gene (for example, samples with a DNM in a recessive gene were excluded). This correction left 584/820 individuals with ID in the RUMC cohort, with 627 DNMs across 584 genes. Similarly, for the ID and neurodevelopmental cohort, we removed all individuals with a DNM in a known ID gene (and other DNMs in these individuals). For the ID cohort, 1,471 samples remained with 1,400 DNMs in 1,235 genes. For the neurodevelopmental cohort, 4,944 samples remained with 4,387 DNM across 3,402 genes (for a complete overview see Supplemental Figure 3.6). We corrected for testing 34,386 genes (i.e., all 18,730 genes minus the 1,537 known ID genes multiplied by 2 for testing the LoF and functional categories).

Validation of the statistical approach by analysis of DNMs in a control cohort

To further confirm the validity of our statistical approach, we applied the same analyses to a set of DNMs identified in trios of healthy individuals and unaffected siblings. For this purpose, we downloaded and reannotated all DNMs identified in 1,911 unaffected siblings

of individuals with ASD from lossifov *et al.*³¹ together with DNMs in controls (Supplementar Table 3.12). In total the control set contained 2,019 coding DNMs found in 2,299 trios. Notably, the protein coding DNM rate in the control cohort was markedly lower than that observed in the individuals with ID (0.91 DNMs versus 1.32 DNMs, respectively). Additionally, we observed no significant enrichment of recurrently mutated genes for LoF or functional mutations ($P = 0.60$ and $P = 0.12$, respectively; Supplemental Figure 3.11).

For the control cohort we performed statistical analyses as described above and identified only one gene that was significantly enriched in functional DNMs. For *YIF1A* (FDR corrected $P = 0.01$) we identified a total of 3 missense DNMs and 1 frameshift DNM (Supplemental Tables 3.13 and 3.14). *YIF1A* may be involved in transport between the endoplasmic reticulum and the Golgi, and it has a pLI of 2.08×10^{-8} indicating this gene is a LoF-tolerant gene. We note that the control cohort consists mostly of healthy siblings from individuals with ASD, and, as such, may still have minor enrichments for mutations that lead to susceptibility to neurodevelopmental disorders.

Increased number of LoF mutations in RUMC cohort compared to controls

To reduce the impact of the enrichment kit used in the control set studies and RUMC cohort, we computed the intersection of all enrichment kits (Agilent SureSelect 37Mb - Agilent SureSelect 50Mb - Agilent SureSelect V4 - Nimblegen SeqCap V2; Supplemental Table 3.12) using the “intersect” function of BEDTools. Only the LoF DNMs present in the 28,189,737-Mb intersection of the four enrichment kits were used in the analysis. The Fisher’s exact test on the enrichment kits normalized LoF DNMs yielded a significant difference with $P = 9.38 \times 10^{-12}$ (RUMC: 157 LoF DNMs of a total of 805 DNMs; controls: 137 LoF DNMs of a total of 1,485 DNMs; OR = 2.38; CI: 1.85–3.07). The coverage and other relevant technical information of the control studies are listed in Supplemental Table 3.12. We note it is important to consider the coverage and false negative rates of all sequencing studies. So far, only a single study has attempted to provide a false negative rate (for example, mutations that are there but were not identified) for exome sequencing, and this was predicted to be <5%.³¹

Attributing pLI for all protein coding genes

To determine the intolerance to LoF variation for each gene, we used the probability of LoF intolerance (pLI), which is based on data from the Exome Aggregation Consortium (ExAC) version 0.3.1, providing exome variants from 60,706 unrelated individuals.⁷¹ The pLI is based on the expected versus observed variant counts to determine the probability that a gene is intolerant to LoF variants and is computed for a total of 18,226 genes. The closer a pLI is to 1, the more intolerant a gene is to LoF variants. The authors consider a $pLI \geq 0.9$ as an extremely LoF-intolerant set of genes. The pLIs for the genes used in this study can be found in Supplemental Table 3.8. Intolerance to LoF variation was evaluated for the available pLIs

of four gene sets: (1) 170 LoF-tolerant (LoFT) genes,¹²² (2) 404 housekeeping genes, involved in crucial roles in cell maintenance,¹²³ (3) 1,359 genes with functional DNMs from the healthy control data set (Supplemental Table 3.14), and (4) 444 well-known dominant ID genes (Supplemental Table 3.4).

Gene set based evaluation of pLI

We evaluated the pLI by computing the expected median pLI for each gene set based on randomly drawing n pLI values from the complete set of 18,226 pLI annotated genes and calculating the median (where n is the number of genes in the gene set). By repeating this random sampling process 100,000 times, we computed the likelihood of the observed median pLI of a gene set compared to the expected median pLI by calculating the empirical P -value:

$$\text{empirical } P = \frac{\left(\sum_{i=1}^N m_i > m_{\text{observed}} \right) + 1}{N + 1}$$

where m is the median pLI of one simulation, m_{observed} is the observed median pLI and N is the total number of performed simulations ($N = 100,000$). In addition, the z -values were computed as described in the section “Evaluating the number of recurrently LoF and functional *de novo* mutated genes.”

Based on the simulations we identified a significantly lower median pLI for the LoFT genes, which is in line with the LoFT nature of this gene set (observed 9.33×10^{-9} distribution simulations: $\mu = 0.04$, $\sigma = 0.03$; empirical $P < 1 \times 10^{-5}$; $z = 1.25$). For the healthy control set, the observed median pLI matched the simulated distribution of median pLI (observed 0.03; distribution simulations: $\mu = 0.03$, $\sigma = 0.01$; empirical $P = 0.31$; $z = 0.39$). For the housekeeping and dominant ID gene sets, the observed median pLI was significantly higher than the simulated distribution of median pLI (HK genes: observed: 0.87; simulated distribution: $\mu = 0.03$, $\sigma = 0.02$; empirical $P < 1 \times 10^{-5}$; $z = 54.05$ and dominant ID genes: observed: 0.95; simulated distribution: $\mu = 0.03$, $\sigma = 0.01$; empirical $P < 1 \times 10^{-5}$; $z = 61.54$). In addition, the median pLI of the housekeeping gene approximated (median pLI = 0.87) and the dominant ID gene set (median pLI = 0.95) surpassed the extremely LoF-intolerant threshold of 0.9, which is in line with the LoF-intolerant nature of housekeeping and dominant ID genes (Supplemental Figure 3.12).

The set of ten novel candidate ID genes has a median pLI of 0.99 (observed 0.99; simulated distribution: $\mu = 0.14$, $\sigma = 0.20$; empirical $P < 1 \times 10^{-5}$; $z = 4.28$) which is, as observed for the dominant ID genes, above the extremely LoF-intolerant gene threshold of 0.9 (Supplemental Figure 3.8). For the missense-only genes (with at least three missense mutations in the absence of LoF mutations, all of which were known dominant ID genes), we observe the

highest median pLI of 0.9999 (observed: 0.9999; simulated distribution: $\mu = 0.09$, $\sigma = 0.14$; empirical $P < 1 \times 10^{-5}$; $z = 6.70$), illustrating that those known and candidate dominant ID genes that harbor only missense variants are among the most LoF-intolerant ID genes (Supplemental Figure 3.8).

Attributing residual variation intolerance scores (RVIS) to all genes

In addition, the residual variation intolerance score (RVIS) was assessed in the same fashion as described for the pLI. The RVIS ranks genes based on whether they have more or less common functional genetic variations relative to the genome-wide expectation. The initial RVIS gene scores were computed based on the NHLBI-ESP6500 data set¹²⁴ and recently recomputed based on the ExAC v0.3 data set (<http://genic-intolerance.org/>). The genes from our study were annotated with RVIS scores based on ExAC (Supplemental Table 3.8).

RVIS scores for gene sets were compared in the same way as for the pLI (Supplemental Figure 3.13). Again, we found the new candidate ID genes to be significantly more intolerant than any random set of genes found (observed median RVIS: 8.47, distribution simulations: $\mu = 50.05$, $\sigma = 15.08$; empirical $P = 4.60 \times 10^{-4}$; $z = -2.76$), like the known dominant ID genes (Supplemental Figure 3.13). For the dominant missense-only genes we again observed the lowest median RVIS of 3.56 (distribution simulations: $\mu = 50.02$, $\sigma = 10.42$; empirical $P < 1 \times 10^{-5}$; $z = -4.46$; Supplemental Figure 3.13).

Estimating clustering of DNMs

The spatial distributions of missense, frameshift and nonsense DNMs were analyzed for clustering within the respective gene they occurred in based on 100,000 simulations. The locations of observed DNMs were randomly sampled over the coding exons of the gene and the distances (in base pairs) between the mutations were normalized for the total coding size of the respective gene. The geometric mean (the n th root of the product of n numbers) of all mutation distances between the DNMs was taken as a measure of clustering. A pseudocount (adding 1 to all distances and 1 to the gene size) was applied to avoid a mean distance of 0 when there were identical mutations.

Based on the prior distance distribution of the 100,000 simulations, we computed a gene-based empirical probability of the observed distance for dominant ID genes with 3 or more DNMs ($n = 64$ genes) in the ID set of 2,104 trios. A total of 21 genes contained only missense mutations (“missense-only” group) and 43 genes contained frameshift, nonsense or a combination of frameshift, nonsense and missense DNMs (“LoF + functional” group). In 21 genes of the missense-only group, 5 genes had empirical probabilities below the significant threshold of 0.05/64, whereas only 1 of the 43 LoF + functional genes had an empirical probability below the significance threshold (Supplemental Table 3.9). Fisher’s exact test was used to compute the statistical significance ($P = 0.012$; OR = 12.56; CI: 1.26–632.65).

Clinical evaluation of selected patients

All patients were referred by clinical geneticists for diagnostic evaluation and overall patient characteristics were comparable to those of a previously published cohort.²⁸ To confirm the identification of the candidate ID genes, we compared the phenotypes of individuals with a DNM in any one of the ten candidate genes and in two genes (*SLC6A1* and *TCF7L2*) significantly enriched in the neurodevelopmental cohort. Comparison of phenotypes was only possible for 8 of 12 genes in which at least 2 individuals with ID were in the RUMC cohort (7 of 10 candidate ID genes, and 1 of 2 candidate NDD genes). A table listing these clinical details is provided in Supplemental Table 3.7. Detailed clinical data for other published individuals is mostly not available. For *TLK2* and *SETD2* a more detailed phenotypic comparison was performed (see Supplemental Note for case studies).

Statistics

Statistical significance was calculated using R statistical computing software version 3.1.0. Two-tailed Fisher's exact tests (significance level α of 0.05) were used to analyze statistical significance between groups for the number of LoF DNMs and number of clustered DNMs. The gene-specific analysis of excess numbers of LoF and functional DNMs was performed using one-sided exact Poisson tests with gene-specific mutation rates taken from the Samocha *et al.* study.¹¹⁵ The gene-specific *P*-values were corrected for multiple testing based on the 18,730 genes present in the Agilent V4 exome enrichment kit times the number of tests (2; LoF and functional), using the Benjamini–Hochberg procedure with an FDR of 0.05. Data distribution was assumed to be normal, but this was not formally tested.

For the statistical testing based on random sampling we used the “sample” and “sample.int” functions (without replacement) from R version 3.1.0 with a random sample size *n* of 100,000. By comparing the observed value to the distribution of the random samples, the empirical *P*-value was computed. In addition the *z*-value was computed by subtracting the mean value of the simulations from the observed value and dividing by the standard deviation of the simulations.



4

***De novo* mutations in SON disrupt RNA splicing of genes essential for brain development and metabolism, causing an intellectual-disability syndrome**

This chapter has been published as:

Jung-Hyun Kim*, Deepali N. Shinde*, **Margot R.F. Reijnders***, Natalie S Hauser, Rebecca L Belmonte, Gregory R Wilson, Daniëlle G. Bosch, Paula A. Bubulya, Vandana Shashi, Slavé Petrovski, Joshua K. Stone, Eun Young Park, Joris A. Veltman, Margje Sinnema, Connie T.R.M. Stumpel, Jos M. Draaisma, Joost Nicolai, University of Washington Center for Mendelian Genomics, Helger G. Yntema, Kristin Lindstrom, Bert BA de Vries, Tamison Jewett, Stephanie L. Santoro, Julie Vogt, The Deciphering Developmental Disorders Study, Kristine K. Bachman, Andrea H. Seeley, Alyson Krokosky, Clessen Turner, Luis Rohena, Maja Hempel, Fanny Kortüm, Davor Lessel, Axel Neu, Tim M. Strom, Dagmar Wiczorek, Nuria Bramswig, Franco A. Laccone, Jana Behunova, Helga Rehder, Christopher T. Gordon, Marlène Rio, Serge Romana, Sha Tang, Dima El-Khechen, Megan T. Cho, Kirsty McWalter, Ganka Douglas, Berivan Baskin, Amber Begtrup, Tara Funari, Kelly Schoch, Alexander P.A. Stegmann, Servi J.C. Stevens, Dong-Er Zhang, David Traver, Xu Yao, Daniel G. MacArthur, Han G. Brunner, Grazia M. Mancini, Richard M. Myers, Laurie B. Owen, Ssang-Taek Lim, David L. Stachura, Lisenka E.L.M. Vissers*, Eun-Young Erin Ahn*

American Journal of Human Genetics (2016) 99, 711-719

* These authors contributed equally

Abstract

The overall understanding of the molecular etiologies of intellectual disability (ID) and developmental delay (DD) is increasing as next generation sequencing technologies identify genetic variants in individuals with such disorders. However, detailed analyses conclusively confirming these variants, as well as the underlying molecular mechanisms explaining the diseases, are often lacking. Here, we report on an ID syndrome caused by de novo heterozygous loss-of-function (LoF) mutations in *SON*. The syndrome is characterized by ID and/or DD, malformations of the cerebral cortex, epilepsy, vision problems, musculoskeletal abnormalities, and congenital malformations. Knockdown of *son* in zebrafish resulted in severe malformation of the spine, brain, and eyes. Importantly, analyses of RNA from affected individuals revealed that genes critical for neuronal migration and cortex organization (*TUBG1*, *FLNA*, *PNKP*, *WDR62*, *PSMD3*, and *HDAC6*) and metabolism (*PCK2*, *PFKL*, *IDH2*, *ACY1*, and *ADA*) are significantly downregulated because of the accumulation of mis-spliced transcripts resulting from erroneous SON-mediated RNA splicing. Our data highlight SON as a master regulator governing neurodevelopment and demonstrate the importance of SON-mediated RNA splicing in human development.

Report

Recent advances in whole-exome and whole-genome sequencing have accelerated the identification of the genetic etiologies of intellectual disability (ID) and developmental delay (DD), facilitating appropriate care and therapy for affected individuals and their families. So far, mutations in more than 1,500 genes have been implicated in ID and DD disorders,^{21; 27-29; 32; 46; 101; 125; 126} and de novo single-nucleotide variants and copy-number variations (CNVs) have been identified as a major cause of severe ID and/or DD.^{28; 46} Recently, two independent studies reported on a single individual with ID and/or DD and a de novo mutation in *SON* (*SON* DNA binding protein [MIM: 182465]), which encodes a protein required for proper RNA splicing. However, the level of evidence required for securely implicating mutations in this gene as disease causing was lacking.^{46; 72; 95} Including these two individuals, we recruited a total of 20 unrelated individuals with mild to severe ID and/or DD (Figure 4.1A and Supplemental Table 4.1) and report on the delineation of an ID syndrome caused by de novo LoF mutations in *SON*. This study was approved by the local institutes under the realm of diagnostic testing. We compared in detail the phenotypic characteristics of all 20 individuals with *SON* LoF mutations. Clinical examination showed that all affected individuals had mild to moderate facial dysmorphisms, including facial asymmetry, midface retraction, low-set ears, downslanting palpebral fissures, deep-set eyes, horizontal eyebrows, a broad and/or depressed nasal bridge, and a short philtrum (Figures 4.1B and Supplemental Figure 4.1). Interestingly, brain MRI, available for 19 affected individuals, revealed that 17 of them had significant abnormalities, including abnormal gyration patterns (polymicrogyria, simplified gyria, and periventricular nodular heterotopia), ventriculomegaly, Arnold Chiari malformations, arachnoid cysts, hypoplasia of the corpus callosum, hypoplasia of the cerebellar hemispheres, and loss of periventricular white matter (Figures 4.1C–E). 11 of 20 individuals developed seizures and/or epilepsy with an age of onset ranging from 1 to 6 years. 17 of 20 individuals presented with musculoskeletal abnormalities, comprising hemivertebrae, scoliosis or kyphosis, contractures, hypotonia, and hypermobility of the joints. Vision problems, including cerebral visual impairment, hypermetropia, optic atrophy, and strabismus, were present in 15 of 20 individuals. In addition, the vast majority of individuals showed congenital malformations consisting of urogenital malformations (6/20), heart defects (5/20), gut malformations (3/20), and a high and/or cleft palate (2/20). Short stature was present in ten individuals, and craniosynostosis involving both the metopic ($n = 1$) and the sagittal sutures ($n = 2$) was noted in 3 of 20 individuals. Metabolic screening was performed in 9 of 20 individuals, confirming mitochondrial dysfunction in individuals 2 and 11 and an O-glycosylation defect in individual 20 (a clinical summary is provided in Table 4.1, and details are listed in Supplemental Table 4.2). Apart from individuals 13 (II-1 in family 13; Figure 1A), 15 (II-3 in family 15), and 20 (II-1 in family 20), none of the individuals had additional coding-sequence mutations that explained (part of) the phenotype (Supplemental Table 4.2). Individual 13 was clinically diagnosed with dyskeratosis congenita, for which a maternally inherited pathogenic *TERT* (MIM: 187270) mutation was identified (Supplemental

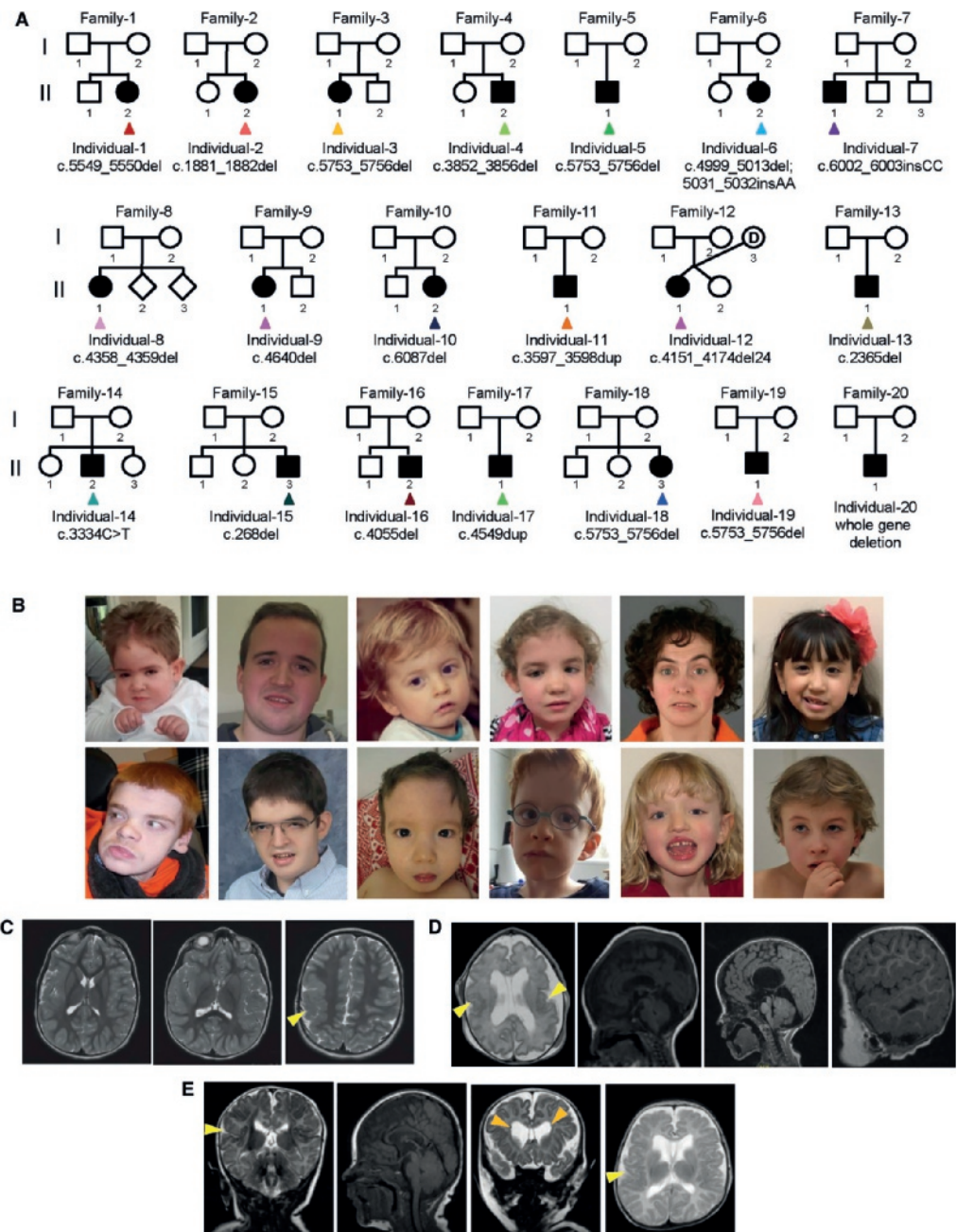


Figure 4.1: Pedigree structures, photos, and brain MRI of individuals with SON mutations
(A) Family pedigrees of individuals carrying mutations in SON. (B) Top row from left to right: photos of individuals 2 (at age 5 years), 4 (age 19 years), 5 (age 2 years), 6 (age 6 years), 8 (age 34 years), and 10 (age 6 years). Bottom row from left to right: photos of individuals 11 (age 21 years), 13 (age 14 years), 15 (age 15 months), 16 (age 5 years), 18 (age 6 years), and 19 (age 10 years). Shared facial dysmorphisms include facial asymmetry, midface retraction, low-set ears, downslanting palpebral fissures, deep-set eyes, horizontal eyebrows, a broad and/or depressed nasal bridge, and a

short philtrum. (C) Axial T2-weighted fast spin-echo MRI of the brain of individual 1 at age 3 years. Three panels show ascending images (left to right) revealing that the individual's insular cortex on the right is thickened and featureless. Less impressive areas of similar change were noted in the posterior aspect of the left insular cortex, which revealed bilateral perisylvian and parietal polymicrogyria (yellow arrowhead). (D) Sagittal T1-weighted and axial T2-weighted MRI of the brain of individual 2. The two images on the left (age 1 day; gestational age 34 + 6 weeks) reveal enlarged lateral ventricles, cavum septum pellucidum, a hypoplastic cerebellar hemisphere, a broad cistern magna, a small fourth ventricle, and a thin corpus callosum. The cortex shows a simplified gyration pattern, and the perisylvian and frontotemporal areas are suspect for polymicrogyria (yellow arrowheads). The two panels on the right (age 2 years) show the fissure Sylvie with an abnormal cortical border, an Arnold Chiari malformation, and hydrocephalus. (E) Frontal T2-weighted, sagittal T1-weighted, axial T2-weighted, and sagittal T1-weighted MRI of the brain of individual 7 (II-1 in family 7) at the age of 2 months. The cortex shows deep sulci and perisylvian areas suspect for polymicrogyria (yellow arrowheads), as well as discrete heterotopic nodules (orange arrowheads). A thin corpus callosum, a small fourth ventricle, enlarged frontal horns of the lateral ventricles, and cavum septum pellucidum are present.

Table 4.2). Individual 13 was, however, more severely affected than could be explained by a *TERT* mutation alone. Similarly, none of the other coding variants identified in individual 15 or the additional genes deleted by the 384 kb deletion CNV in individual 20 were likely to explain the phenotype of these individuals (Supplemental Table 4.2). *SON* (GenBank: NM_138927.2) is composed of 12 exons (Figure 4.2A) and encodes a protein (GenBank: NP_620305.2) containing an arginine/serine (RS)-rich domain and two RNA-binding motifs (a G-patch and a double-stranded RNA binding motif) (Figure 4.2A).¹²⁷⁻¹²⁹ 17 of 20 mutations are frameshift mutations, including a recurrent 4-bp deletion (c.5753_5756del) in four independent individuals (Supplemental Table 4.1 and Figures 4.2A and 4.2B). The remaining mutations include a nonsense mutation, an in-frame deletion of eight amino acids, and a complete gene deletion (Supplemental Table 4.1). Importantly, parental DNA was available for testing in 19 of 20 individuals and indicated that all mutations had occurred de novo (Supplemental Figure 4.2 and Supplemental Table 4.1). Interestingly, de novo truncating mutations in *SON* have not been observed in over 2,000 control individuals,^{29; 31; 40; 130; 131} and *SON*, with a Residual Variation Intolerance Score of 1.88, belongs to the 2% most intolerant human protein-coding genes.¹²⁴ Furthermore, interrogation of large databases (such as the Exome Aggregation Consortium [ExAC] Browser) has shown that, after sequence context and mutability are considered, *SON* is significantly depleted of LoF variants according to multiple LoF metrics (pLI = 1.00, and the false-discovery rate of the LoF depletion score is $p = 1.68 \times 10^{-6}$).^{71; 132} Although these population genetic signatures of intolerance cannot be considered sufficient evidence of causality on their own, they support the hypothesis that *SON* LoF mutations are under strong purifying selection in the human population and that their occurrence most likely contributes to severe clinical phenotypes. Transcripts bearing a premature stop codon are likely to be degraded by nonsense-mediated mRNA decay. To confirm that LoF mutations result in reduced dosage of *SON*, we used three different PCR primer sets (Supplemental Table 4.3) to perform qRT-PCR to determine the amounts of the *SON* transcript in peripheral-blood mononuclear cells (PBMCs) isolated from trio 1 (I-1, I-2, and II-2 in family 1), trio 3 (I-1, I-2, and II-1 in family 3), individual 5 (II-1 in family 5; Figure 4.1A), and an unrelated healthy donor (Figure 4.2C). All three primer sets showed that compared to mRNA from the parental samples and the unrelated healthy donor, *SON* mRNA

Table 4.1 Clinical features of individuals with SON haploinsufficiency

	Percentage	Number of affected individuals
Intellectual disability	100%	20/20
Brain malformation	89%	17/19
Ventricular enlargement	74%	14/19
Corpus callosum abnormality	53%	10/19
Cortex malformation	37%	7/19
White matter abnormalities	21%	4/19
Cerebellum abnormalities	21%	4/19
Other	11%	2/19
Neurological features	85%	17/20
Seizures	55%	11/20
Hypotonia	75%	15/20
Musculoskeletal abnormalities	85%	17/20
Hypermobility	40%	8/20
Scoliosis or kyphosis	20%	4/20
Hemivertebrae	10%	2/20
Contractures	10%	2/20
Other	85%	17/20
Eye and/or vision abnormality	75%	15/20
Strabismus	55%	11/20
Suspicion of CVI	20%	4/20
Hypermetropia	30%	6/20
Heart defect	25%	5/20
Gastrointestinal malformation	15%	3/20
Urogenital malformation	30%	6/20
Horseshoe kidney	10%	2/20
Other	20%	4/20
Facial dysmorphism	100%	20/20
Short stature	50%	10/20
Craniosynostosis	15%	3/20

Abbreviation: CVI = cortical visual impairment

in the affected individuals was significantly downregulated (Figure 4.2C). Subsequent western blotting using PBMC lysates from trio 1 and two different SON antibodies consistently showed the reduction of SON in individual 1 (Figures 4.2D and 4.2E), indicating that de novo SON LoF mutations result in SON haploinsufficiency. To examine the effect of SON haploinsufficiency on embryonic development, we utilized *Danio rerio* (zebrafish), which

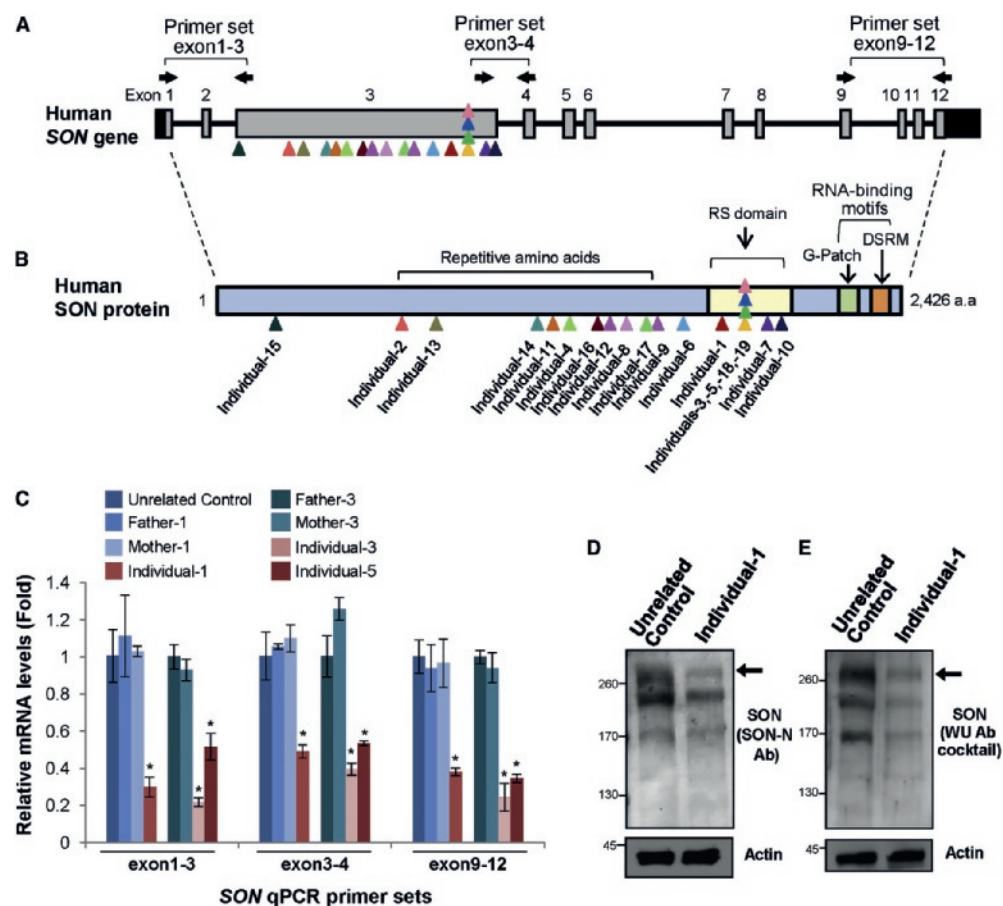


Figure 4.2: SON mutations and their functional effect at the RNA and protein levels.

(A and B) Schematic representation of SON (A) and SON (B) shows the position of the mutations identified in the 20 affected individuals with color-coded arrowheads. The locations of the PCR primer sets are indicated by black arrows. (C) Real-time qPCR with three different primer pairs showed that SON mRNA from the affected individuals was overall downregulated in comparison to mRNA from the parents and unrelated normal individual. Error bars represent mean \pm SD. * $p < 0.001$. (D and E) Western blotting demonstrated reduced expression of SON. SON-N antibody (1:1,000) was generated against amino acids 74–88 of the human SON (amino acid sequence DTELRYKPDLEKGSR). The cocktail of WU SON antibodies was a mixture of three different SON antibodies (WU09 [1:100], WU14 [1:2,000], and WU21 [1:200]). The epitopes of WU SON antibodies were as follows: MDSQMLATSS for WU09, CEESESKTKSH for WU14, and SMMSAYERS for WU21. SON-N antibody (D) and the cocktail of SON WU antibodies (E) showed similar results. The bands indicated by the black arrow represent full-length SON. Other bands, which could represent potential isoforms, were also detected. Besides the bands present in samples from both normal and affected individuals, no other specific bands were detected in the affected individuals.

has a well-conserved homolog of human SON (NCBI Gene: LOC565999; Supplemental Figures 4.3 and 4.4). We assessed the developmental effects of *SON* haploinsufficiency in vivo with morpholino (MO)-mediated knockdown of *son* in zebrafish embryos. Interestingly, embryos injected with a *son* MO had a host of developmental defects that ranged from bent tails (63.6%) to eye malformations and microcephaly (17.1%) and shortened or gnarled tails, deformed body axes, and massive body curvatures (2.1%) 24 hr post-injection (hpi) (Figure 4.3A and Supplemental Figure 4.5). Embryos that survived 72 hpi progressed to more severe phenotypes including extreme spinal malformations (22.2%), head and eye malformations with brain edema (37.2%), and profound developmental abnormalities (10.1%; Figure 4.3B), mimicking features observed in the affected individuals. SON is a nuclear speckle protein able to bind to both DNA and RNA, and its cellular functions include regulation of RNA splicing and gene transcription, as well as proper cell-cycle and embryonic stem cell maintenance.^{128; 133-136} To identify molecular mechanisms underlying the clinical features of individuals with *SON* haploinsufficiency, we examined global expression patterns upon *SON* knockdown in cellular systems. Hereto, we re-analyzed microarray-based RNA-expression profiling and RNA-sequencing datasets generated upon *SON* knockdown in HeLa cells^{128; 133} and human embryonic stem cells.¹³⁶ Surprisingly, from these previous datasets, we noticed

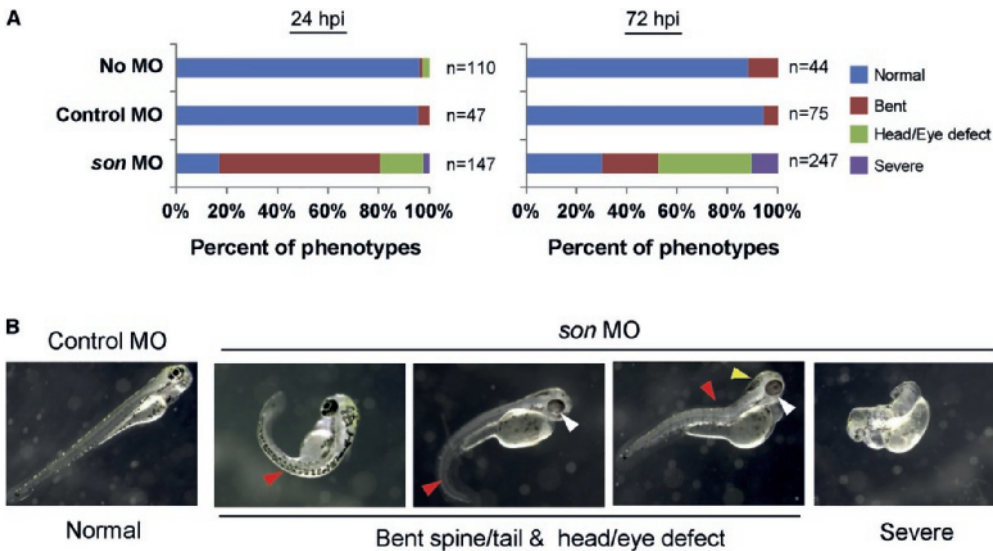


Figure 4.3: Targeted *son* knockdown in developing zebrafish causes impaired head development and spinal malformations. (A) Zebrafish injected with a splice-blocking *son* morpholino (MO; 5'-TGGTCCTGGGATATAACAGACAGATT-3', 6.25 ng) that targeted the junction between intron 9 and exon 10, a control MO (5'-CCTCTTACCTCAGTTACAATTTATA-3', 6.25 ng), or no MO showed a normal phenotype, a bent spine or tail, a head or eye defect, or a severe phenotype at 24 and 72 hpi. The percentages of embryos with each phenotype are shown in the bar graphs, and the number of embryos examined is listed next to each bar. (B) Representative images of the phenotype observed 72 hr after MO injection (red arrow, bent spine or tail; white arrow, eye defects; and yellow arrow, brain edema).

that a group of genes playing pivotal roles in neuronal cell migration, embryonic survival, metabolism, and mitochondrial function, including *TUBG1* (MIM: 191135), *FLNA* (MIM: 300017), *PNKP* (MIM: 605610), *WDR62* (MIM: 613583), *PSMD3*, *HDAC6* (MIM: 300272), *PCK2* (MIM: 614095), *PFKL* (MIM: 171860), *IDH2* (MIM: 147650), *ACY1* (MIM: 104620), and *ADA* (MIM: 608958), showed significantly decreased expression upon *SON* knockdown (Supplemental Tables 4.4 and 4.5).^{128; 133; 136} To investigate whether genes involved in regulating brain development and in metabolism are also downregulated in individuals with *SON* LoF mutations, we measured the levels of RNA expression of these genes in PBMCs from trio 1, trio 3, and individual 5, as well as from an unrelated healthy donor (primers are listed in Supplemental Table 4.3). Using qPCR analysis, we confirmed that all 11 genes were indeed significantly downregulated in individuals with *SON* haploinsufficiency (Figures 4.4A and 4.4B). *SON* functions as a splicing co-factor that promotes correct and efficient RNA splicing of weak splice sites and alternative splice sites by facilitating spliceosome recruitment to the elongating RNA polymerase II complex.¹²⁸ Prominent features observed upon *SON* knockdown in HeLa cells and human embryonic stem cells have included intron retention and exon skipping, which have been shown at the gene level for *TUBG1*, *HDAC6*, and *ADA*.^{128; 133; 136} We next sought to determine whether RNA splicing of these 11 genes is also impaired in our individuals with *SON* haploinsufficiency. To this end, we analyzed the pre-mRNA sequences of the remaining eight genes to predict weak splice sites that could be potential targets of *SON*-mediated RNA splicing (Supplemental Table 4.6). We performed RT-PCR by using DNase-treated RNA samples isolated from trio 1, trio 3, individual 5, and an unrelated healthy donor and using primers designed to detect intron retention and exon skipping (Supplemental Tables 4.6 and 4.7). We not only confirmed that these sites were indeed mis-spliced in HeLa cells upon *SON* knockdown (Supplemental Figure 4.6) but also found that all three affected individuals showed significant intron retention (*TUBG1*, *FLNA*, *PNKP*, *WDR62*, *PSMD3*, *PCK2*, *PFKL*, *IDH2*, and *ACY1*) and exon skipping (*HDAC6* and *ADA*) at the predicted sites of the target pre-mRNAs and that this resulted in the accumulation of mis-spliced products (Figures 4.4C and 4.4D). In contrast, mis-spliced RNA products were absent in the parents and unrelated donor (Figures 4.4C and 4.4D). Together, these results indicate that *SON*-mediated RNA splicing is severely compromised in individuals with *SON* haploinsufficiency. Our data have revealed that the complex neurodevelopmental disorder observed in these affected individuals is due to compromised *SON* function, which causes insufficient production of downstream targets as a result of erroneous *SON*-mediated RNA splicing. Moreover, the roles of several downregulated genes are well-known causes of ID and/or DD in humans (Supplemental Tables 4.4 and 4.5).^{29; 32; 137-146} For instance, *FLNA* haploinsufficiency is the most common cause of periventricular nodular heterotopia (MIM: 300049),¹⁴⁶ a rare brain malformation that we also found among our cohort with *SON* LoF mutations. Similarly, de novo LoF mutations in *TUBG1* are known to result in cortical malformations (MIM: 615412),¹⁴³ also frequently observed in our cohort of affected individuals. Because we have shown that a substantial number of essential

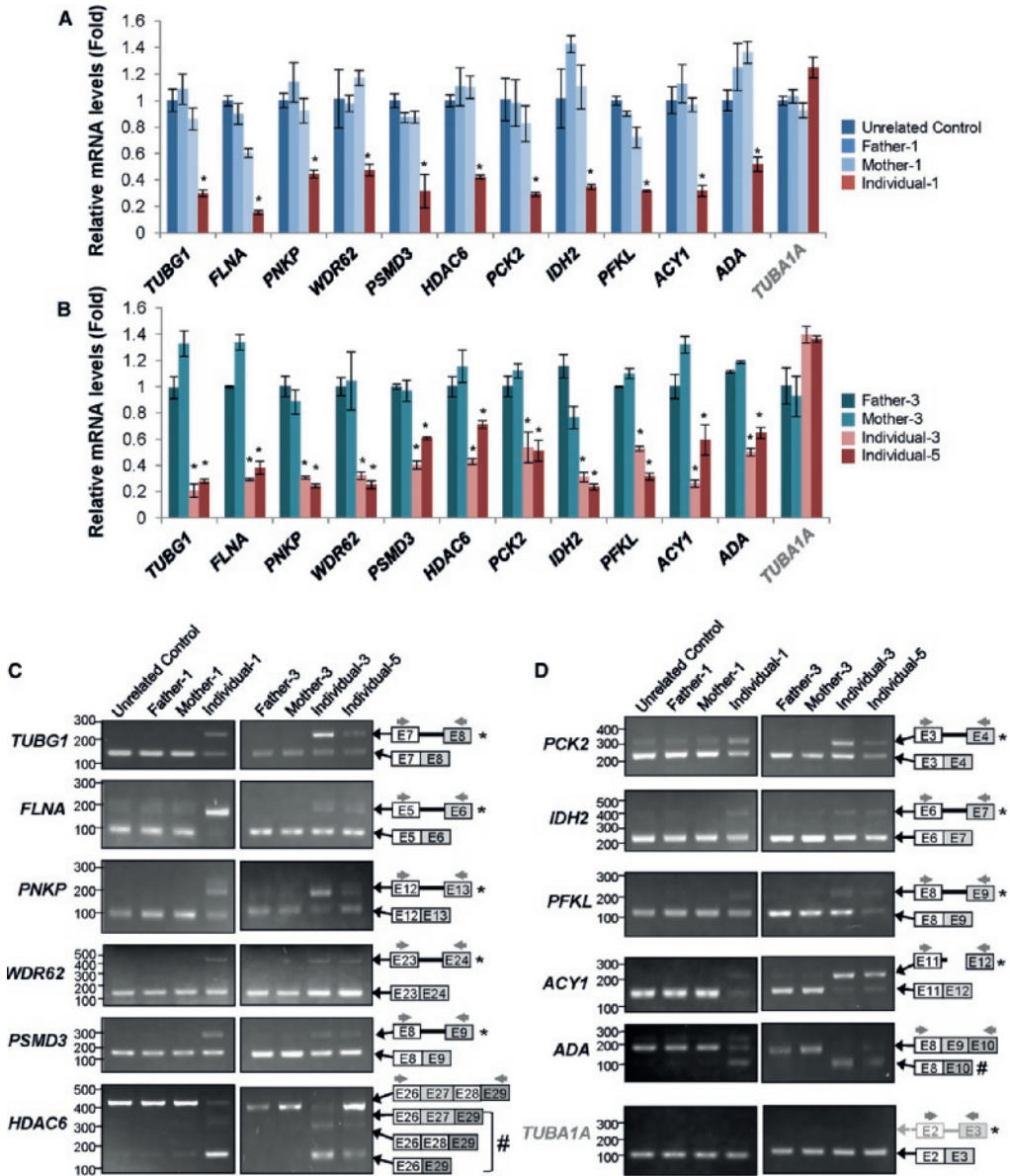


Figure 4.4: Individuals carrying heterozygous SON LoF mutations have defective RNA splicing of genes associated with the pathophysiology of ID and/or DD and metabolic disorders, resulting in their reduced expression.

(A and B) Multiple genes associated with the pathophysiology of ID and/or DD (A) and metabolic disorders (B) in the affected individuals were downregulated in comparison to genes from parents and unrelated healthy individuals. TUBA1A mRNA served as a negative control (unaffected transcript). Error bars represent mean ± SD. **p* < 0.001. (C and D) Intron retention and exon skipping of genes involved in ID and/or DD when mutated (C) and genes involved in metabolic disorders when mutated (D) in the individuals with SON mutations. The locations of the primers used for PCR are marked by gray arrows above the exons. Analysis of TUBA1A pre-mRNA, which served as a negative control, demonstrated that splicing of this transcript is not impaired in the affected individuals. *, intron-retained products; #, exon-skipped products.

developmental genes are significantly downregulated upon *SON* haploinsufficiency, *SON* thus represents a master regulator of genes essential for human neurodevelopmental processes. In summary, we have identified de novo LoF mutations in *SON* as a cause of a complex neurodevelopmental disorder associated with ID and/or DD and severe brain malformations. In addition, we have revealed the underlying molecular mechanism by showing that *SON* haploinsufficiency leads to defective RNA splicing of multiple genes critical for brain development, neuronal migration, and metabolism. Our findings thus greatly contribute to our understanding of how defective RNA splicing leads to human neurodevelopmental disorders.



5

***De novo* and inherited loss-of-function variants in *TLK2*: identification, clinical delineation and genotype-phenotype evaluation of a novel neurodevelopmental disorder.**

A revised version of this chapter has been published as:

Margot R.F. Reijnders*, Kerry A. Miller*, Mohsan Alvi, Jacqueline A.C. Goos, Melissa M. Lees, Anna De Burca, Alex Henderson, Alison Kraus, Barbara Mikat, Bert B.A. de Vries, Bertrand Isidor, Bronwyn Kerr, Carlo Marcelis, Caroline Schluth-Bolard, Charu Deshpande, Claudia A.L. Ruivenkamp, Dagmar Wiczorek, The Deciphering Developmental Disorders Study, Diana Baralle, Edward M. Blair, Hartmut Engels, Hermann-Josef Lüdecke, Jacqueline Eason, Gijs W.E. Santen, Jill Clayton-Smith, Kate Chandler, Katrina Tatton-Brown, Katelyn Payne, Katherine Helbig, Kelly Radtke, Kimberly M. Nugent, Kirsten Cremer, Lauren Brick, Lynne M. Bird, Margje Sinnema, Maria Bitner-Glindzicz, Marieke F. van Dooren, Marielle Alders, Marije Koopmans, Mariya Kozenko, Megan L. Harline, Merel Klaassens, Michelle Steinraths, Nicola S. Cooper, Patrick Edery, Patrick Yap, Paulien A. Terhal, Peter J. van der Spek, Phillis Lakeman, Rachel L. Taylor, Rebecca O. Littlejohn, Rolph Pfundt, Saadet Mercimek-Andrews, Alexander P.A. Stegmann, Sarina G. Kant, Scott McLean, Shelagh Joss, Sigrid M.A. Swagemakers, Sofia Douzgou, Steven A. Wall, Sébastien Küry, Eduardo Calpena, Nils Koelling, Simon J. McGowan, Stephen R.F. Twigg, Irene M.J. Mathijssen, Christoffer Nellaker, Han G. Brunner*, Andrew O.M. Wilkie*

American Journal of Human Genetics (2018) 102, 1195-1203

* These authors contributed equally

Abstract

Next generation sequencing is a powerful tool for the discovery of novel genes related to neurodevelopmental disorders (NDDs). Here, we report the identification of a new syndrome due to *de novo* or inherited heterozygous mutations in the Tousled-like kinase 2 gene (*TLK2*) in 38 unrelated individuals and two affected mothers using whole exome and whole genome sequencing technologies, matchmaker databases and international collaborations. Affected individuals had a consistent phenotype, characterized by mild-borderline neurodevelopmental delay (86%), behavioral disorders (68%), severe gastrointestinal problems (63%), and facial dysmorphism including blepharophimosis (82%), telecanthus (74%), prominent nasal bridge (68%), broad nasal tip (66%), thin vermilion of the upper lip (62%), and upslanting palpebral fissures (55%). Analysis of cell lines from three affected individuals showed that mutations act through a loss-of-function mechanism in at least two cases. Genotype-phenotype analysis and comparison of computationally modeled faces showed that phenotypes of these and other individuals with loss-of-function variants significantly overlapped with phenotypes of individuals with other variant types (missense and C-terminal truncating). This suggests that haploinsufficiency of *TLK2* is the most likely underlying disease mechanism, leading to a consistent neurodevelopmental phenotype. This work illustrates the power of international data sharing, by the identification of 40 individuals from 26 different centers in seven different countries, allowing the identification, clinical delineation and genotype-phenotype evaluation of a novel NDD caused by mutations in *TLK2*.

Report

The introduction of whole exome sequencing (WES) as a diagnostic test for individuals with unexplained neurodevelopmental disorders (NDDs) has led to the identification of dozens of novel disease genes. As a recent example, statistical analysis of aggregated exome data uncovered variants in ten different genes as likely causes of intellectual disability, a subtype of NDDs characterized by deficits in both intellectual and adaptive functioning.^{1; 126} One such gene was Tousled-like kinase 2 (*TLK2*), which was originally named because of homology to the *Arabidopsis* gene *Tousled*.¹⁴⁷ The *TLK2* gene, ubiquitously expressed in all tissues including fetal brain, encodes a serine/threonine kinase comprising a catalytic domain and multiple highly conserved coiled-coil motifs.^{147; 148} *TLK2* is known to have maximal activity during the S-phase of the cell cycle and is therefore tightly linked to DNA-replication.¹⁴⁷ DNA double-strand breaks lead to rapid and transient inhibition of *TLK* activity, suggesting a role in DNA repair.¹⁴⁹ With the discovery of both H3-H4 chaperone Asf1 and histone H3 as physiological substrates of *TLKs*, its protein function has been linked to chromatin assembly.¹⁵⁰⁻¹⁵⁴

To establish the contribution of *TLK2* variants to NDDs in humans, we systematically collected phenotypic data of the five affected individuals with *TLK2* variants reported previously,¹²⁶ derived cell lines and exploited different strategies to identify additional individuals with a variant in this gene. By using Genematcher⁸⁸ and by sharing data with international collaborators, we identified a total of 38 unrelated individuals and two affected mothers with heterozygous variants in *TLK2*. Variants were detected by either family-based WES (research settings: $n=18$ probands and 2 affected parents; diagnostic settings: $n=18$ probands) or whole genome sequencing (WGS) (research settings, $n=2$ probands) in 26 different institutions and seven different countries (Supplemental Figure 5.1; Supplemental Methods, Supplemental Tables 5.1 and 5.2). Two additional individuals with *de novo* *TLK2* variants p.(Val505Asp) and p.(Arg724Gln), each of whom had a second likely pathogenic mutation in another gene, were excluded from further consideration to avoid confounding in the phenotypic analysis (Supplemental Methods). IRB-approved consents for WES or WGS in diagnostic or research settings were obtained for all individuals.

We observed a broad spectrum of different variant types in *TLK2* (NM_006852): four frameshift variants, ten nonsense variants (including two located in the last exon), twelve canonical splice-site variants, and nine missense variants (Figures 5.1A-1C; Table 5.1). Additionally, we identified a *de novo* balanced translocation in one of the WGS cases, resulting in a breakpoint at chromosome 17 disrupting the *TLK2* intron between exons 2 and 3 (Figure 5.1D; Supplemental Methods). Interestingly, we found recurrent mutations within our cohort of affected individuals. Missense variants p.(His496Arg) and p.(Arg339Trp) were each reported in two unrelated individuals. In addition, p.(Arg339Gln) affects the same amino acid residue as variants p.(Arg339Trp) (Figure 5.1C; Table 5.1). Finally, two splice variants

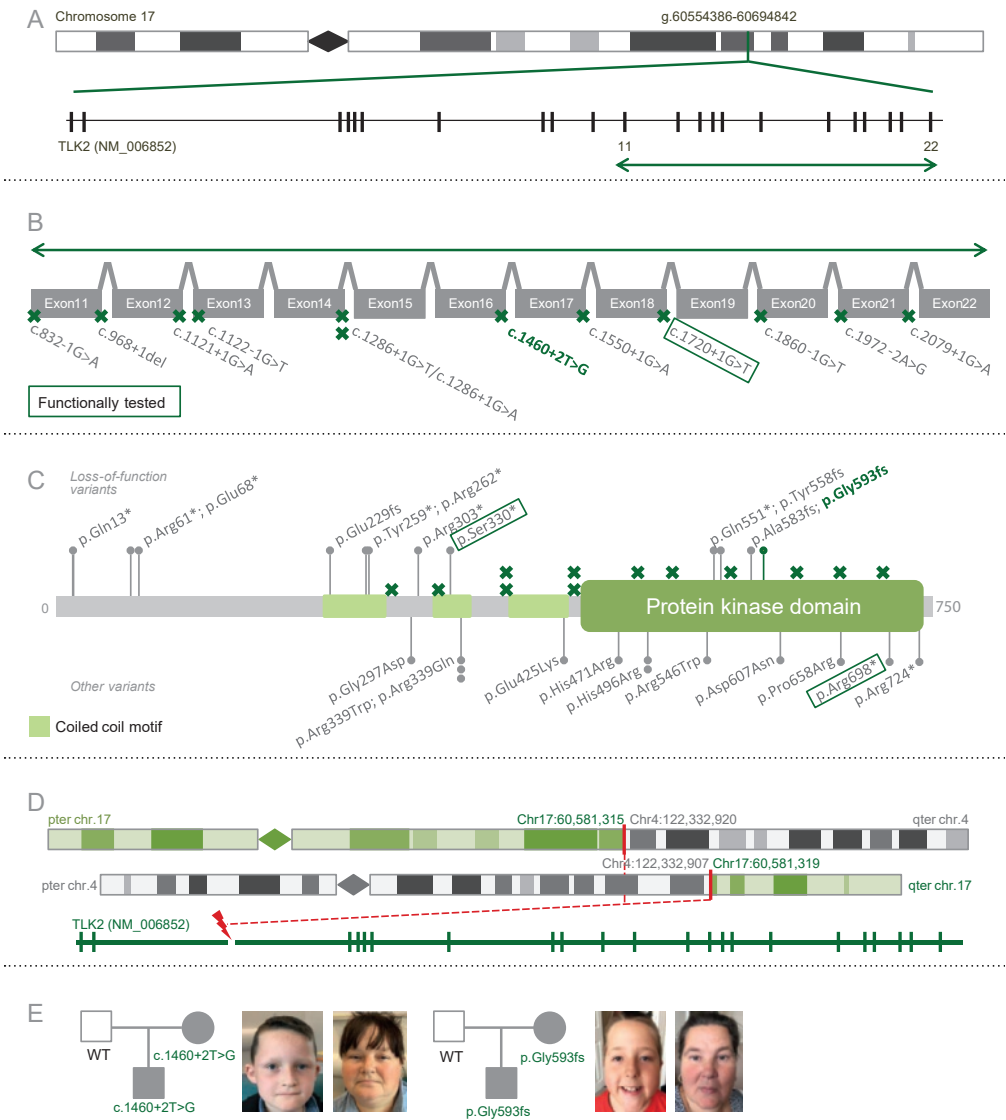


Figure 5.1: Intragenic variants and balanced translocation identified in TLK2

(A) Location of TLK2 (NM_006852) in chromosome 17 (see Supplemental Information for discussion about different TLK2 spliceforms). Vertical marks in TLK2 represent the 22 exons. Green arrow indicates region enlarged in panel below. (B) Schematic view (not to scale) of exons 12-22 and location of identified splice site mutations (green crosses). The splice site mutation inherited from an affected parent is shown in bold and green. The variant subjected to cDNA analysis is shown in the dark green rectangle. (C) Overview of TLK2 protein with the protein kinase domain (dark green) and three coiled coil motifs (light green). Loss-of-function variants (13; 8 nonsense and 5 frameshift) are shown above the protein and other variants (23; 10 splice site variants, shown as green crosses, 11 missense variants and 2 nonsense variants causing a premature stop codon in the last exon). The frameshift mutation inherited from an affected parent is shown in bold and green. The variants subjected to cDNA analysis are shown in the dark green rectangle (D) Balanced translocation between chromosomes 4 and 17, with the breakpoint disrupting TLK2 between exons 2 and 3, identified in one individual:46,XX,t(4;17)(27;q23.2).seq[GRCh37]t(4;17)g.[chr4:pter_cen_122332907::chr17:60,581,319_qter]g.[chr17_pter_cen_60,581,315::chr4:122,332,920_qter] (E) Pedigrees of individuals with inherited variants and photographs of probands and their affected mothers. Both mothers have facial dysmorphism similar to their child.

were predicted to give rise to the same affected protein product: c.1286+1G>T, p.(?) and 1286+1G>A, p.(?) (Figure 5.1B; Table 5.1). From the nine missense variants identified in eleven unrelated individuals, five are located in the catalytic domain of the protein and three in a coiled coil motif. One variant, p.(Gly297Asp), is located outside a known functional domain, but affects a highly conserved amino acid and was predicted pathogenic by several *in silico* prediction programs, similar to other missense variants (Figure 5.1C; Supplemental Table 5.3). None of the missense variants were present in the ExAC database,⁷¹ nor in our in-house database of variants identified in healthy controls. The recently released gnomAD database, containing WGS variants identified in controls, reported only p.(Arg546Trp) in a single individual (allele frequency of ~0.000004). None of the other missense variants were present in the gnomAD database (Table 5.1).

For all but two variants (Table 5.1), the *de novo* status was assessed by sequencing the parents of the proband. In two individuals, variants were inherited from a similarly affected parent, while all other variants ($n=34$) occurred *de novo*. Detailed phenotyping revealed that both mothers carrying a predicted loss-of-function (LOF) *TLK2* variant (Table 5.1) were mildly affected. The first mother (c.1460+2T>G) had mild neurodevelopmental delay and speech delay. The second affected mother (p.Gly593fs) had low-normal IQ levels, but was diagnosed with bipolar disorder. Both had facial dysmorphism similar to their affected child (Figure 5.1E). The inherited mutations illustrate that the search for a diagnosis should not always be restricted to *de novo* mutations, in particular if individuals are only mildly affected. Similar to the parents in this study, who were never referred for genetic testing before diagnostic testing in their child uncovered a *TLK2* mutation, we expect mutations causing milder phenotypes to be present in the general population. This could explain why, although *TLK2* exhibits very strong constraint against LOF variants ($pLI = 1$), five LOF variants (low-coverage variants excluded) have been reported in gnomAD, and a missense variant p.(Arg546Trp) reported here as *de novo* variant, was present at very low allele frequency in the population (aggregate minor allele frequency of LOF and missense variants ~0.000024).

Consistent with the phenotypes of both affected mothers, mild neurodevelopmental phenotypes accompanied by language and motor delay, were present in the majority of the 38 unrelated probands: 6% of the individuals had normal IQ levels (85-100), 14% had borderline ID (IQ 70-85) and from the 72% diagnosed with ID (IQ<70), most had mild ID (IQ 50-70) (Figure 5.2). Most of the affected probands (22 males and 16 females) were children at the time of last examination (median 8.0 years; interquartile range 4.1-13.5 years), ages ranged between 3 months and 29 years. Three individuals, who all had language and motor delay, were too young for formal assessment of their neurodevelopmental phenotype. In addition to this, systematic evaluation of other clinical data, scored by the referring clinician, showed a variety of overlapping features (Figure 5.2, Supplemental Table 4). Neurological problems including hypotonia (37%), epilepsy (13%), and non-specific intracranial brain

Table 5.1 identified intragenic variants in TLK2 (NM_006852), inheritance and presence in ExAC and gnomAD databases*

Subgroup	cDNA position	Protein position	Inheritance	RNA analysis	cMAF Exac	cMAF gnomAD
Predicted LOF	c.37C>T	p.(Gln113*)	<i>De novo</i>	No	No LOF variants	5 LOF variants: ~0.00002
	c.181C>T	p.(Arg61*)	<i>De novo</i>	No		
	c.202G>T	p.(Glu68*)	<i>De novo</i>	No		
	c.685_688del	p.(Glu229fs)	<i>De novo</i>	No		
	c.777C>A	p.(Tyr259*)	<i>De novo</i>	No		
	c.784C>T	p.(Arg262*)	<i>De novo</i>	No		
	c.832-1G>A	p.(?)	<i>De novo</i>	No		
	c.907C>T	p.(Arg303*)	<i>De novo</i>	No		
	c.968+1del	p.(?)	<i>De novo</i>	No		
	c.989C>A	p.(Ser330*)	<i>De novo</i>	Yes		
	c.1121+1G>A	p.(?)	<i>De novo</i>	No		
	c.1122-1G>T	p.(?)	<i>De novo</i>	No		
	c.1286+1G>T	p.(?)	<i>De novo</i>	No		
	c.1286+1G>A	p.(?)	<i>De novo</i>	No		
	c.1460+2T>G	p.(?)	Inherited	No		
	c.1550+1G>A	p.(?)	<i>De novo</i>	No		
	c.1651C>T	p.(Gln551*)	<i>De novo</i>	No		
	c.1672dup	p.(Tyr558Leufs*4)	<i>De novo</i>	No		
	c.1720+1G>T§	p.(?)	<i>De novo</i>	Yes		
	c.1746delA	p.(Ala583Argfs*5)	<i>De novo</i>	No		
	c.1776_1783delTGGTCTTT	p.(Gly593Glufs*5)	Inherited	No		
	c.1860-1G>T	p.(?)	Unknown	No		
	c.1972-2A>G	p.(?)	<i>De novo</i>	No		
	c.2079+1G>A	p.(?)	<i>De novo</i>	No		
Other variant types	c.2092C>T§	p.(Arg698*)	<i>De novo</i>	Yes	0	0
	c.2170C>T	p.(Arg724*)	<i>De novo</i>	No	0	0
	c.890G>A	p.(Gly297Asp)	<i>De novo</i>	No	0	0
	c.1015C>T	p.(Arg339Trp)	<i>De novo</i> ∞	No	0	0
	c.1016G>A	p.(Arg339Gln)	<i>De novo</i>	No	0	0
	c.1273G>A	p.(Glu425Lys)	Unknown	No	0	0
	c.1412A>G§	p.(His471Arg)	<i>De novo</i>	No	0	0
	c.1487A>G§	p.(His496Arg)	<i>De novo</i> ∞	No	0	0
	c.1636C>T	p.(Arg546Trp)	<i>De novo</i>	No	0	~0.000004
	c.1819G>A§	p.(Asp607Asn)	<i>De novo</i>	No	0	0
	c.1973C>G	p.(Pro658Arg)	<i>De novo</i>	No	0	0

Abbreviations: cMAF = cumulative minor allele frequency; LOF = Loss-of-function

* Identified balanced translocation (n=1) is not included in this table

§ Variant reported previously¹²⁶

∞ Recurrent de novo variant identified in two unrelated individuals

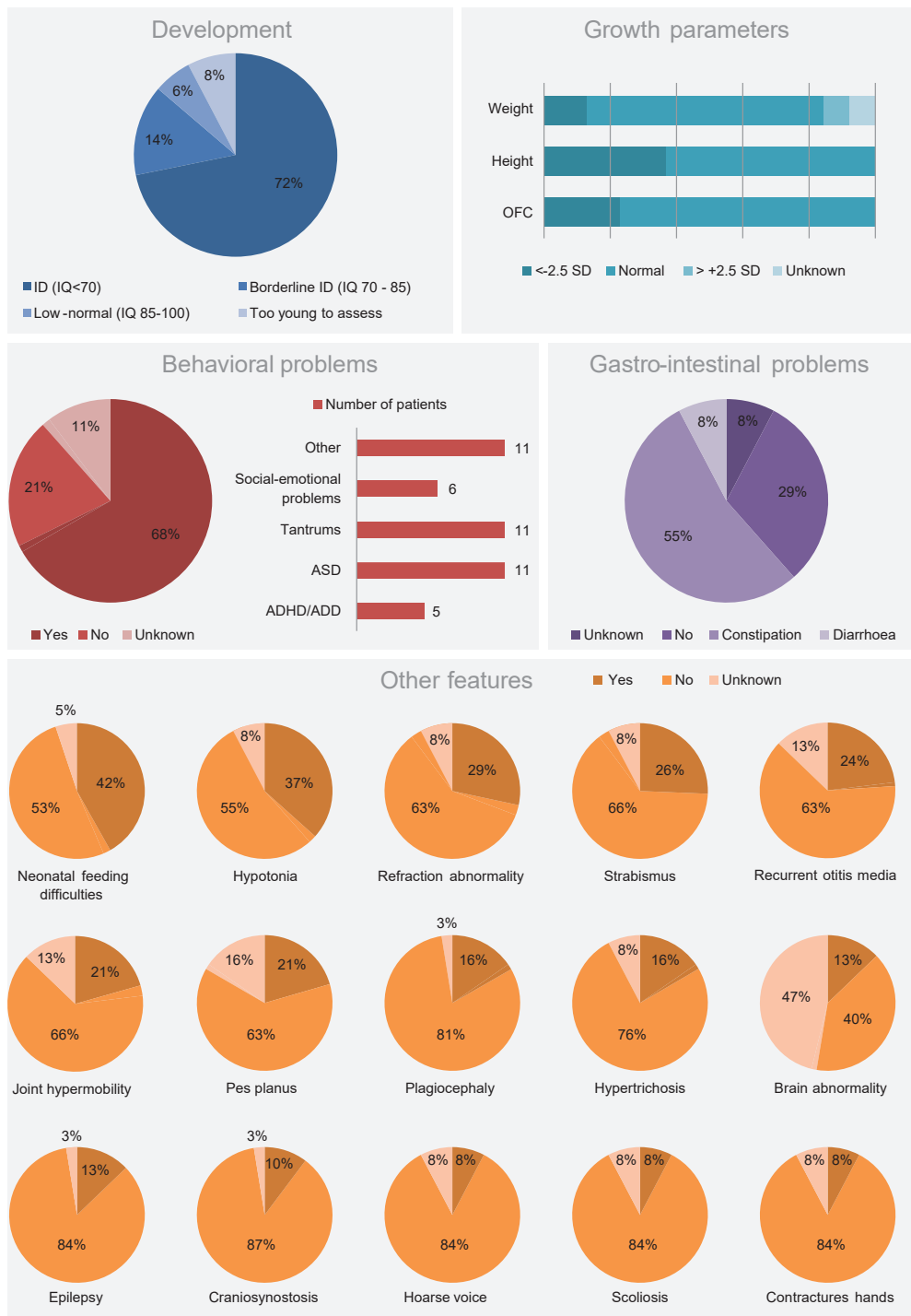


Figure 5.2: Clinical spectrum associated with *TLK2* variants
Overview of clinical features observed in individuals with *TLK2* variants.

abnormalities (13%) (Supplemental Table 5) were observed. A broad range of behavioral disorders was present (68%), with often severely affected social functioning: tantrums (11 individuals), autism spectrum disorder (11 individuals), attention-deficit disorder with or without hyperactivity (5 individuals) and severe social-emotional problems (6 individuals) were the most commonly reported problems. Less frequently observed were short attention span, pica disorder, aggression, obsessive-compulsive disorder and anxiety in eleven individuals. Other recurrent features included gastro-intestinal problems (constipation in 55%; severe diarrhea in 8%), neonatal feeding difficulties (42%), eye abnormalities (refraction abnormality in 29%, strabismus in 26%), musculoskeletal abnormalities (joint hypermobility in 21%; pes planus in 21%; toe walking in 18%; scoliosis in 8%; contractures of the hands in 8%), recurrent otitis media (24%), hypertrichosis (16%) and hoarse voice (8%). Abnormalities of skull shape were observed in 31% of probands (Figure 5.2, Supplemental Tables 4 and 6), with clinically proven craniosynostosis being present in four (10%) of them (Supplemental Table 7). However, sequence-based screening of 309 DNA samples from individuals with mixed, genetically undiagnosed craniosynostosis did not identify further cases, indicating that *TLK2* mutations are only rarely associated with craniosynostosis. Growth parameters were frequently abnormal. Short stature was documented in 37%, microcephaly in 24% and low body weight in 13%. Three individuals (8%) were overweight. Features reported in only one or two individuals are summarized in Supplemental Table 6. In addition to the other clinical features, overlapping facial dysmorphisms were present (Figures 5.3A and B). Most frequently reported by clinicians were blepharophimosis (82%), telecanthus (74%), prominent nasal bridge (68%), broad nasal tip (66%), thin vermilion of the upper lip (62%), and upslanting palpebral fissures (55%). Pointed and tall chin (42%), epicanthal folds (42%), narrow mouth (32%), high palate (30%), microtia (29%), posteriorly rotated ears (29%), long face (27%), ptosis (21%), and asymmetric face (16%) were observed in fewer than half of the individuals.

Analysis of data from the ExAC database demonstrates that *TLK2* is extremely intolerant for LOF variants (pLI score = 1).⁷¹ In line with this observation, animal models with depletion of *TLK2* have been reported to have severely disturbed cellular and developmental processes. *Drosophila* with complete LOF of *TLK* were associated with arrested nuclear divisions,

Figure 5.3: Facial dysmorphism of individuals with *TLK2* variants

(A) Photographs of 21 unrelated individuals with a loss-of-function variant in *TLK2*, showing overlapping facial dysmorphism. Most frequently reported by clinicians were blepharophimosis, telecanthus, prominent nasal bridge, broad nasal tip, thin vermilion upper lip, and upward slanted palpebral fissures. Pointed and tall chin, epicanthal folds, narrow mouth, high palate, microtia, posteriorly rotated ears, long face, ptosis, and asymmetric face were observed in less than half of the individuals. (B) Photographs of 7 unrelated individuals with a missense or C-terminal truncating variant in *TLK2*. Variant p.(Arg724*) is assigned to this subgroup, since a premature stop codon is introduced in the last exon. Facial dysmorphisms overlapped with dysmorphism observed in individuals with loss-of-function variants. (C) Computational averaging of 33 facial photographs of 22 subjects with LOF variants in *TLK2* (left) compared with 22 gender- and age-matched controls (right). (D) Computational averaging of 11 facial photographs of 8 subjects with missense or C-terminal truncating variants in *TLK2* (left) compared with 8 gender- and age-matched controls (right).

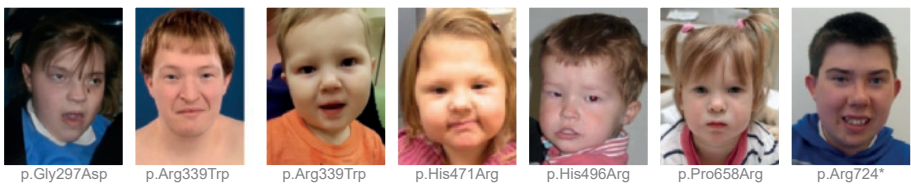
A

PATIENTS WITH LOSS-OF-FUNCTION VARIANTS

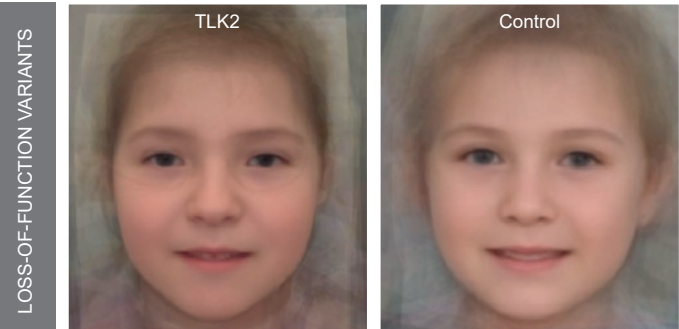


B

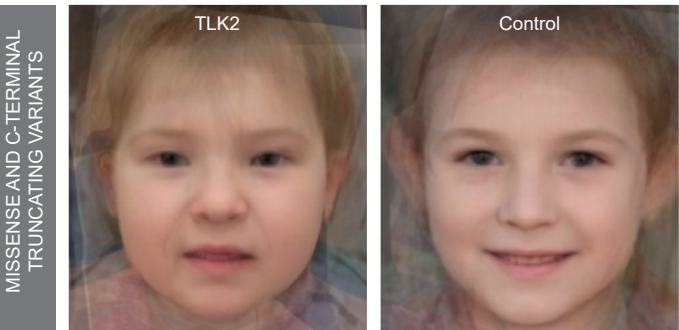
PATIENTS WITH MISSENSE AND C-TERMINAL TRUNCATING VARIANTS



C



D



causing apoptosis of the cell.¹⁵¹ *Tlk2* null mice were embryonically lethal due to placental failure.¹⁵⁵ In this study, we found several predicted LOF variants in affected individuals. To investigate whether variants resulted in an aberrant transcript, we synthesized cDNA from RNA extracted from fibroblast or lymphoblastoid cell lines from three individuals with different variants: (1) p.(Ser330*), predicted to result in a truncated product leading to nonsense-mediated decay (NMD); (2) p.(Arg698*), with a premature stop codon in the last exon predicted to escape from NMD; and (3) c.1720+1G>T, a mutation predicted to affect splicing of exon 18. To investigate the significance of NMD for expression of *TLK2* transcripts, we treated fibroblasts (for p.Ser330*) and lymphoblastoid cell lines (for p.(Ser330*), p.(Arg698*) and c.1720+1G>T) with cycloheximide, an inhibitor of NMD.¹⁵⁶ Transcript stability of cDNA PCR products from p.(Ser330*) and p.(Arg698*) individuals in the presence of cycloheximide was analyzed using a restriction enzyme assay targeting the wild-type transcript and the results were confirmed using deep sequencing to quantify relative levels of wild-type and mutant transcripts (Supplemental Methods). For fibroblast and lymphoblastoid cell lines heterozygous for the p.(Ser330*) variant, the mutant allele represented 15.8% and 21.5% of transcripts respectively in the absence of cycloheximide, but rose to 37.7% and 48.5% respectively in the presence of cycloheximide, supporting that this variant is subject to NMD and causes haploinsufficiency of *TLK2*. In contrast, wild-type and mutant transcripts from lymphoblastoid cells of the individual heterozygous for p.(Arg698*) did not show significant differences between treated and untreated cells, supporting that the mutant transcript escapes NMD due to its location within the last coding exon of *TLK2* (Figure 5.4A). Amplification of cDNA from an individual with a splice-site variant (c.1720+1G>T) showed a full length wild-type product of 300 bp and an additional aberrant smaller product of 130 bp, consistent with skipping of exon 18. Direct sequencing of this smaller fragment confirmed that exon 17 spliced directly to exon 19, thereby producing an out-of-frame transcript predicted to introduce a premature stop codon at the next amino acid position (p.Ser517fs*1). Additionally, the intensity of the spliced transcript increased when treated with cycloheximide, indicating that the mutant transcript is subjected to NMD (Figure 5.4B).

By analyzing *TLK2* transcripts in cell lines of three different individuals, we were able to confirm that transcripts were subjected to NMD in two of them, causing haploinsufficiency of *TLK2*. It is likely that comparable variants predicted to cause LOF of *TLK2* affect the transcript similarly. The large number of identified individuals with *TLK2* variants allowed us to search for underlying pathogenic mechanisms for the individuals with variants with unknown effect, such as p.(Arg698*). To assess this, we divided our cohort in two subgroups and (1) performed a structured genotype-phenotype analysis, and (2) created and compared computationally modeled faces. Subgroup 1 ($n=25$) included all probands carrying a predicted LOF variant (nonsense, frameshift variant or canonical splice-site, or balanced translocation) similar to variants p.(Ser330*) and c.1720+1G>T. Subgroup 2 ($n=13$) comprised individuals with either missense variants or variants causing a premature stop codon in the

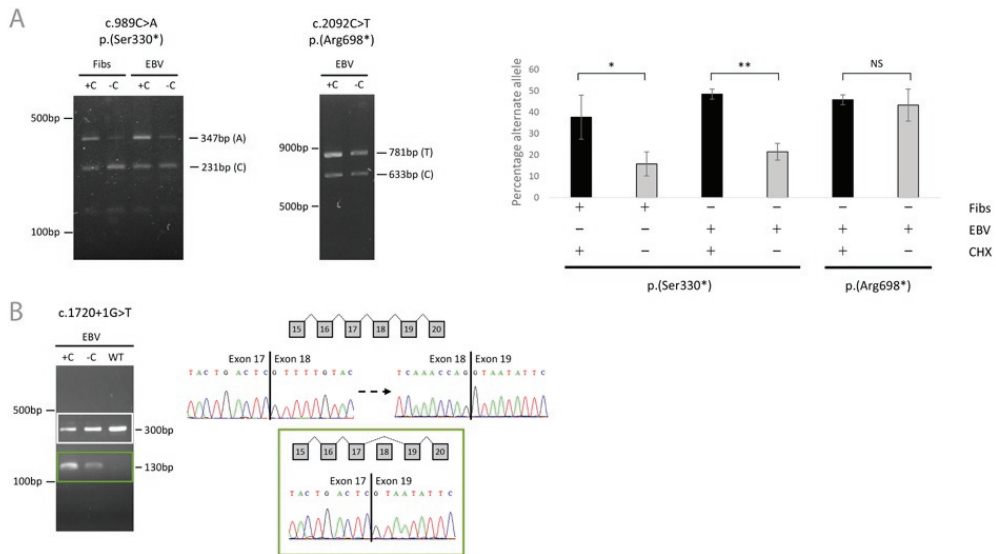


Figure 5.4: Analysis of transcripts of three different cell lines

(A) Analysis of transcripts encoding nonsense mutations p.(Ser330*) and p.(Arg698*) in cell lines of affected individuals. Left panel shows reverse transcriptase-PCR (RT-PCR) products of cDNA prepared from fibroblast and lymphoblastoid cell lines of subject with p.(Ser330*) variant, either in the presence (+C) or absence (-C) of cycloheximide and incubated with Apol (digests wild-type allele). Central panel shows RT-PCR of cDNA prepared from lymphoblastoid cell line of subject with p.(Arg698*) variant, in the presence (+C) or absence (-C) of cycloheximide and incubated with Hpy99I (digests wild-type allele). Right panel shows proportion (\pm standard deviation) of variant alleles quantified by deep sequencing of triplicate samples. Statistical testing of differences: *, $P=0.046$; **, $P=0.011$; NS, not significant.

(B) Analysis of transcripts with canonical splice-site mutation c.1720+1G>T. A wild-type fragment at 300 bp in c.1720+1G>T lymphoblastoid cells is observed as well as a second fragment at 130 bp, which is absent in control cDNA. An increase of mutant transcript in cells was present when treated with cycloheximide (+C), indicating that the aberrant transcript was subject to NMD. Sequencing of the 300 bp (white box) and 130 bp (green box) fragments demonstrated skipping of exon 18 in the lower cDNA product. Abbreviations: Fibs = fibroblasts; EBV = lymphoblastoid cells; C/CHX = cycloheximide; WT = control cDNA.

last exon of *TLK2*, such as p.(Arg698*). Affected parents of probands with inherited mutations were not included in the subgroups. Next, we compared frequencies of 40 different features and frequencies of 15 facial dysmorphisms between the two groups using a two-tailed Fisher's exact test. This showed that both clinical features and facial dysmorphisms were remarkably similar between the two subgroups. From the 55 different features, none differed significantly between the two subgroups ($p < 0.05$), even without correction for multiple testing (Supplemental Table 4). Secondly, averaged visualization of facial dysmorphism by computational modeling of 33 photographs from 22 individuals in subgroup 1 and 11 photographs from 8 individuals in subgroup 2 at different ages (Supplemental Methods), showed consistent differences from a comparable number of gender- and age-matched controls, including telecanthus, broad nasal tip and tall, pointed chin (Figure 5.3C and D).

Given this strong overlap in phenotypes and facial dysmorphic features between probands with different type of mutations, it is likely that not only LOF variants, but also the majority of identified missense variants and variants with a premature stop codon in the last exon only have a single functional copy of *TLK2*. Hence, we conclude that the predominant pathogenic mechanism of these *TLK2* mutations is haploinsufficiency.

Often mentioned together with *TLK2* is its close interactor *TLK1*. From birth, murine *Tlk2* shows a similar expression pattern to the closely related paralog *Tlk1* across many tissues.¹⁵⁵ Human *TLK1* has 84% identity to *TLK2* at the protein level,¹⁴⁷ and it was shown that *TLK1*-depletion leads to extensive chromosome segregation defects in human cells.¹⁵⁷ Interestingly, *TLK1* is (similarly to *TLK2*) intolerant for both missense and truncating mutations in healthy individuals (significant z-scores of 3.84 (*TLK1*) and 5.67 (*TLK2*) and pLI (constraint) scores of 1.00 for both *TLK1* and *TLK2*) (ExAC database).⁷¹ In the literature, four *de novo* variants have been reported in *TLK1* (NM_012290.4): p.(Pro25Leu) in an individual with intellectual disability,¹²⁶ p.(Met566Thr) in an individual with autism,¹⁵⁸ p.(Ala599Gly) in an individual with a NDD and congenital heart disease¹⁵⁹ and p.(Lys367fs) in an individual with schizophrenia.³⁹ Importantly, none of these variants are present in the ExAC or gnomAD databases. Taking this into account, it is possible that *TLK1* variants could contribute to NDDs, similar to the homolog *TLK2*. In future research, the exact role of *TLK1* in NDDs should be further explored.

In conclusion, we show that both *de novo* and inherited mutations in *TLK2* cause a novel neurodevelopmental disorder, hallmarked by mild developmental delay, a variety of behavioral disorders, severe gastro-intestinal problems and facial dysmorphism. The identification of a large number of individuals ($n=40$, including two affected mothers) emphasizes the power and importance of data sharing, allowing us to delineate the clinical phenotype and to evaluate genotype-phenotype correlations. By analyzing three cell lines of affected individuals, we were able to confirm that at least two variants act through a heterozygous loss-of-function mechanism (haploinsufficiency). Phenotypes of these individuals and others with comparable loss-of-function variants significantly overlapped with phenotypes of individuals with other variant types, providing further evidence for the underlying disease mechanism of the *TLK2* variants. Given the genetic and functional similarities between *TLK2* and *TLK1*, further research should focus on the potential role of *TLK1* mutations in developmental disorders.



6

***RAC1* mutations in developmental disorders with diverse phenotypes**

This chapter has been published as:

Margot R. F. Reijnders*, Nurhuda M. Ansor*, Maria Kousi*, Wyatt W. Yue, Perciliz L. Tan, Katie Clarkson, Jill Clayton-Smith, Ken Corning, Julie R. Jones, Wayne W.K. Lam, Grazia M.S. Mancini, Carlo Marcelis, Shehla Mohammed, Rolph Pfundt, Maian Roifman, Ronald Cohn, David Chitayat, Deciphering Developmental Disorders Study, Tom H. Millard, Nicholas Katsanis, Han G. Brunner*, Siddharth Banka*

American Journal of Human Genetics (2017) 101, 466-477

* These authors contributed equally

Abstract

RAC1 is a widely studied Rho GTPase, a class of molecules that modulate numerous cellular functions essential for normal development. RAC1 is highly conserved across species and is under strict mutational constraint. We report seven individuals with distinct *de novo* missense *RAC1* mutations and varying degrees of developmental delay, brain malformations, and additional phenotypes. Four individuals, each harboring one of c.53G>A (p.Cys18Tyr), c.116A>G (p.Asn39Ser), c.218C>T (p.Pro73Leu), and c.470G>A (p.Cys157Tyr) variants, were microcephalic, with head circumferences between -2.5 to -5 SD. In contrast, two individuals with c.151G>A (p.Val51Met) and c.151G>C (p.Val51Leu) alleles were macrocephalic with head circumferences of +4.16 and +4.5 SD. One individual harboring a c.190T>G (p.Tyr64Asp) allele had head circumference in the normal range. Collectively, we observed an extraordinary spread of ~10 SD of head circumferences orchestrated by distinct mutations in the same gene. *In silico* modeling, mouse fibroblasts spreading assays, and *in vivo* overexpression assays using zebrafish as a surrogate model demonstrated that the p.Cys18Tyr and p.Asn39Ser *RAC1* variants function as dominant-negative alleles and result in microcephaly, reduced neuronal proliferation, and cerebellar abnormalities *in vivo*. Conversely, the p.Tyr64Asp substitution is constitutively active. The remaining mutations are probably weakly dominant negative or their effects are context dependent. These findings highlight the importance of *RAC1* in neuronal development. Along with *TRIO* and *HACE1*, a sub-category of rare developmental disorders is emerging with *RAC1* as the central player. We show that ultra-rare disorders caused by private, non-recurrent missense mutations that result in varying phenotypes are challenging to dissect, but can be delineated through focused international collaboration.

Report

Developmental disorders (DDs) are etiologically extremely heterogeneous and affect 2%–5% of individuals.^{20; 21} *De novo* mutations account for a substantial proportion of DDs and are thought to underlie approximately 400,000 new DD-affected case subjects world-wide annually.³³ Recently, large-scale next generation sequencing studies have led to the identification of several DD-associated genes that lead to clinical manifestations through protein truncating or recurrent missense variants.^{31; 32; 126} However, rare disorders caused by private, non-recurrent missense mutations that result in varying phenotypes remain challenging to dissect.⁷²

Several human DDs are known to result from mutations in members of the RAS superfamily of small GTPases.¹⁶⁰ The RAS superfamily is further divided into smaller families, one of which is the 22-member Rho family. Rho GTPases cycle between active GTP-bound and inactive GDP-bound states. Their activity is regulated by guanine nucleotide exchange factors (GEFs), GTPases activating proteins (GAPs), and guanine nucleotide dissociated inhibitors (GDIs).¹⁶¹ Rho GTPases modulate essential cellular functions, including cell polarity, migration, vesicle trafficking, and cytokinesis and play crucial roles in neuronal development, neuronal survival, and neurodegeneration.^{162–164} However, no human DDs caused by mutations in genes encoding Rho GTPases are known.

One of the most widely studied Rho GTPases is the RAS-related C3 Botulinum Toxin Substrate 1 (RAC1).¹⁶⁵ RAC1 is part of the RAC Rho GTPases subfamily that also includes RAC2, RAC3, and RhoG.¹⁶⁶ RAC1 is an important modulator of the cytoskeleton, with a critical function in phagocytosis, mesenchymal-like migration, neuronal polarization, axonal growth, adhesion, and differentiation of multiple cell types.^{162; 167; 168} Additionally, it is involved in cellular growth and cell-cycle regulation via mTOR signaling.¹⁶⁹ In mouse studies, Rac1 is required for the formation of three germ layers during gastrulation, with *Rac1*-knockout mice being embryonic lethal.¹⁷⁰ Conditional forebrain-specific *Rac1*-knockout mice display impaired neuronal migration, abnormal dendritic growth and remodelling, disruption of lamellipodia formation, reduced neuronal proliferation, premature differentiation, and microcephaly.^{171–173} Here, we report *de novo* missense *RAC1* (MIM: 602048) mutations in individuals with DD and divergent phenotypes.

All procedures followed were in accordance with the ethical standards of the institutional and national responsible committees on human or animal experimentation and that, where relevant, informed consent was obtained. Review of data from 4,293 families, who underwent trio whole-exome sequencing (WES) as part of the Deciphering Developmental Disorders study,³³ led to identification of three individuals with *de novo* *RAC1* (GenBank: NM_006908) missense mutations: individual 3 with c.218C>T (p.Pro73Leu), individual 5 with c.190T>G

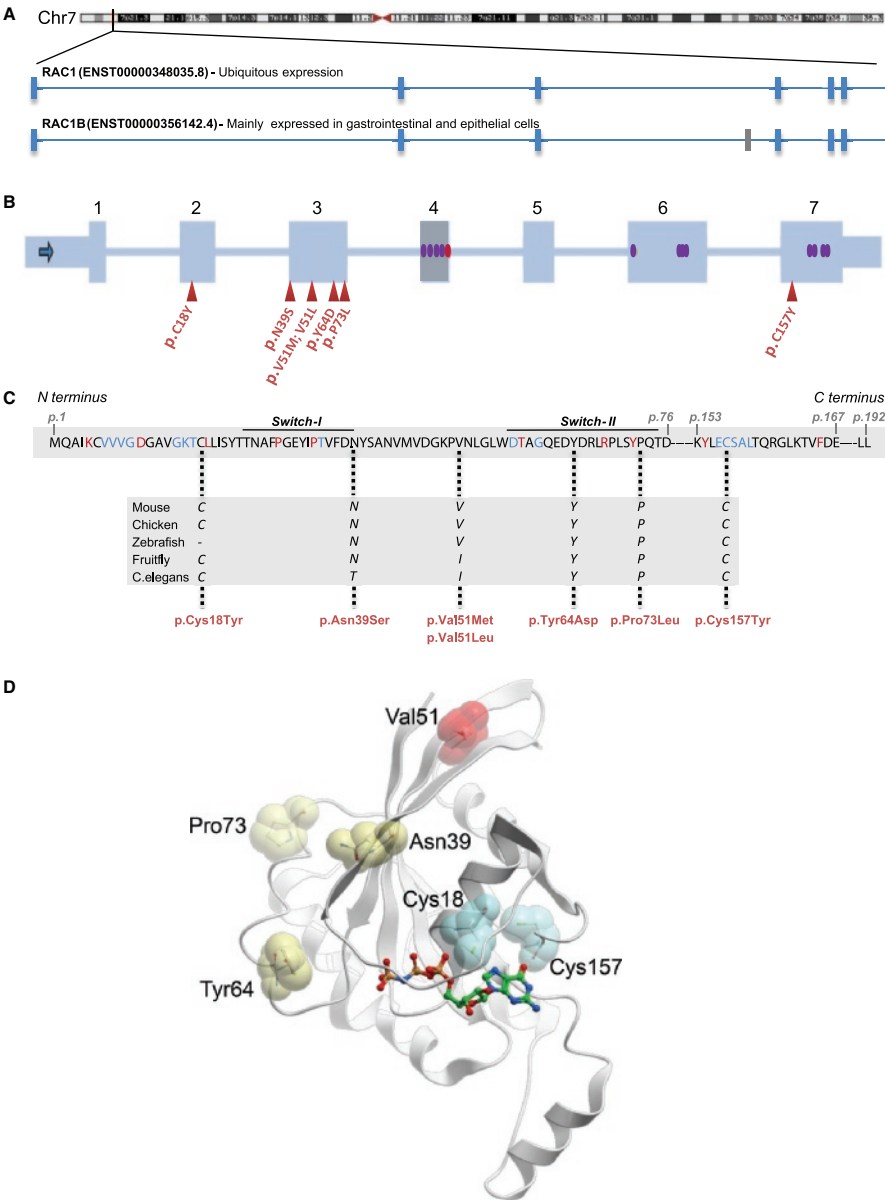


Figure 6.1: Properties of germline RAC1 variants

(A) RAC1 is located on human chromosome 7p22.1. Exon-intron structure of the two protein-coding RAC1 transcripts is shown. The ubiquitously expressed RAC1 transcript is composed of six exons (blue vertical lines). RAC1B (mainly expressed in gastrointestinal and epithelial cells) contains an additional exon (exon 4; gray vertical line). (B) Schematic of RAC1, with the coding exons shown as blue rectangles and the introns as blue lines. Exon 4 that is expressed only in the longer RAC1B transcript is shown as a gray rectangle. The red triangles represent the positions of the mutations identified in this study. The red oval within exon 4 shows the two splice donor variants detected in the ExAC database. The purple ovals represent the missense variants in the ExAC database. (C) Amino acid sequence of the RAC1 across positions p.1-p.192. Red amino acids are highly conserved in >90% of the RAS family members and blue amino acids represent RAS protein G box consensus residues. Positions of the identified missense mutations (in red) are marked

with dotted lines and the high conservation among different species of these amino acids is shown in the rectangle below. (D) Crystal structure of human RAC1 showing the position of the identified missense variants presented in this study. The sites of substitution are indicated by sticks in spheres and are color coded according to their putative impact on guanine nucleotide binding (cyan), protein-protein interactions (yellow), and structural stability (red). For reference, the non-hydrolyzable GTP analog GMPPNP bound to the structure is shown in sticks. The structure was analyzed and this figure was generated using the program ICM-Pro v3.8 (Molsoft, LLC).

(p.Tyr64Asp), and individual 6 with c.151G>A (p.Val51Met) (Table 6.1; Figures 6.1 and 6.2). Two additional individuals were independently ascertained through family-based diagnostic WES: individual 1 with c.53G>A (p.Cys18Tyr) and individual 2 with c.116A>G (p.Asn39Ser) (Table 6.1; Figures 6.1 and 6.2). While functional studies were ongoing for the five individuals, two additional individuals were identified via the GeneMatcher tool⁸⁸ or through international collaboration: individual 4 with c.470G>A (p.Cys157Tyr) and individual 7 with c.151G>C (p.Val51Leu) (Table 6.1; Figures 6.1 and 6.2).

Human *RAC1* (ENSG00000136238) encodes six transcripts, of which two are protein coding: RAC1 and RAC1B (Figure 6.1A). Of the two protein-coding transcripts, GenBank: NM_006908 (RAC1, ENST00000348035.8) lacks exon 4 and encodes the shorter RAC1 isoform of 192 amino acids, which is ubiquitously expressed in all tissues.¹⁷⁴ *RAC1* is under strict mutational constraint with only 15 missense variants observed versus 75.9 expected in ~60,000 exomes cataloged in the ExAC database (z-score = 3.42).⁷¹ Moreover, 7 of these 15 observed missense variants are located in exon 4, which is included only in RAC1B (ENST00000356142.4) that encodes the longer isoform and is mainly expressed in gastrointestinal and epithelial tissues.¹⁷⁵ Of particular note, there are no missense variants in exons 1, 2, 3, and 5 in the ExAC database (Figure 6.1B). Six out of seven mutations described here are in exons 2 and 3. The p.Cys157Tyr change (p.Cys176Tyr in the longer RAC1B) lies in a sub-region of exon 7 with no known germline human missense variants (Figure 6.1B). None of the identified mutations were present in any of the in-house variant databases of the four centers participating in this study. All six amino acids affected by the seven mutations are highly conserved among different species (Figure 6.1C). All reported *RAC1* mutations are within or in proximity to G box residues and/or conserved residues present in 90% of the RAS superfamily members (Figure 6.1C).¹⁷⁶ Collectively, the genetic data were strongly supportive of deleteriousness for each of the seven mutations.

Informed consent for publication of photographs was obtained from legal guardians. Detailed clinical information was collected on all affected individuals (Table 6.1; Figure 6.2; see Supplemental Note). All seven individuals (age range 4.5 months–15 years) had moderate to severe intellectual disability (ID) and variable degrees of neurological involvement including hypotonia (4/7), epilepsy (3/7), behavioral problems (3/7), and stereotypic movements (2/7). However, their occipital frontal circumferences (OFC) were remarkably different: individuals 1, 2, 3, and 4 were microcephalic (OFCs of –2.5, –3, –5, and –2.5 SD, respectively), individual

Table 6.1 *RAC1* mutations and phenotypes

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7
	Microcephaly	Microcephaly	Microcephaly	Microcephaly	Normal OFC	Macrocephaly	Macrocephaly
Gender	Male	Male	Male	Male	Male	Male	Male
Ethnicity	Caucasian/Egypt	Caucasian	Caucasian	Caucasian/Armenian	Caucasian/Asian	Caucasian	African American
Age of examination	13 years	9 years	15 years	4.5 months	12 years	33 months	4 years 5 months
Mutation (NIM_006908)							
Chromosome position (hg19)	Chr7:6426860G>A	Chr7:6431563A>G	Chr7:6431665C>T	chr7:6441968G>A	Chr7:643163T>G	Chr7:6431598G>A	Chr7:6431598G>C
cDNA change	c.53G>A	c.116A>G	c.218C>T	c.470G>A	c.190T>G	c.151G>A	c.151G>C
Amino acid change	p.(Cys18Tyr)	p.(Asn39Ser)	p.(Pro73Leu)	p.(Cys157Tyr)	p.(Tyr64Asp)	p.(Val51Met)	p.(Val51Leu)
Likely effect of the mutation	Dominant negative	Dominant negative	Unknown	Unknown	Constitutively active	Unknown	Unknown
Growth							
Height	127 cm (-2.5 SD)	128 cm (-2.5 SD)	Unknown	62 cm (-1 SD)	134 cm (+1.03 SD) (8 yrs)	Unknown	109 cm (0 SD)
Weight	30 kg (0 SD)	24 kg (-0.5 SD)	Unknown	4.8 kg (-3 SD)	Unknown	Unknown	23 kg (+2.5 SD)
Head circumference	50 cm (-2.5 SD)	47.7 cm (-3 SD)	47 cm (-5 SD)	39 cm (-2.5 SD)	56.5 cm (+1 SD)	57 cm (+4.16 SD)	59.5 cm (+4.5 SD)
Development							
Intellectual disability	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mild/Moderate/Severe	Moderate	Mild-Moderate	Severe	Unknown	Severe	Moderate	Unknown
Neurological							
Epilepsy	Yes	No	Unknown	Yes	No	Unknown	Yes
Hypotonia	Yes	No	Unknown	Yes	Yes	Unknown	Yes
Behavioural problems	No	Yes - Hyperactive	Unknown	Unknown	Yes - Sleep disturbances	Unknown	Yes - Autism

(Continues on the next page)

Stereotypic movements	No	No	Unknown	No	Yes	Unknown	Yes
Brain MRI abnormalities							
Cerebellar abnormalities	Yes	Yes	Unknown	Yes	No	No	No
Hypoplasia corpus callosum	Yes	Yes	Unknown	Yes	Yes	No	No
Enlarged lateral ventricles	Yes	Yes	Unknown	No	No	No	No
Enlarged fourth ventricle	No	Yes	Unknown	No	No	No	No
Thin pons brain stem	No	Yes	Unknown	Yes	No	No	No
Mega cisterna magna	Yes	Yes	Unknown	Yes	No	No	No
Polymicrogyria	No	No	Unknown	No	Yes	No	No
White matter lesions	Yes	No	Unknown	No	No	Yes	Yes
Congenital abnormalities							
Cardiac abnormalities	Yes – NS LVC; IV	No	Unknown	Yes – PDA, PFO, BAV	Yes – VSD	Unknown	No
Hypospadia	No	No	Unknown	Yes	Yes	Unknown	No
Other							
Neonatal feeding difficulties	Yes	Yes	Yes	No	No	Unknown	No
Other	Plagiocephaly; Scoliosis; Small hands and feet; hyperlaxity; brachydactyly 5th digit; SC bilateral	Recurrent pneumonias; Eczema	Diabetes Mellitus	umbilical hernia; tracheobronchomalacia; cryptorchidism	Mild visual impairment; congenital sensorineural hearing impairment; Abnormal creases hand	None	Non-verbal: Recurrent otitis media; Eczema

Abbreviations: BAV = bicuspid aortic valve; IV = insufficiency all valves; NS LVC = Non-synchronous left ventricle contractions; PDA = patent ductus arteriosus; PFO = patent foramen ovale; SC = simian crease; VSD = ventricular septum defect

5 with p.Tyr64Asp mutation had a normal OFC (+1 SD), and individuals 6 and 7, both with mutations affecting Val51, were macrocephalic (OFCs of +4.16 and +4.5 SD, respectively) (Table 6.1; Figure 6.2). Hypoplasia of the corpus callosum and the cerebellar vermis were the commonest features observed on available magnetic resonance imaging (MRI) studies of individuals with microcephaly. However, they all had additional abnormalities (Supplemental Note; Figure 6.2). Individual 5 (with normal OFC) was reported to have polymicrogyria and hypoplastic corpus callosum (images not available). The two individuals with macrocephaly (individuals 6 and 7) showed periventricular white matter lesions (Table 6.1; Figure 6.2). Arched eyebrows, dysplastic ears, prominent nasal bridges, and overhanging columellae were shared between a majority of individuals without macrocephaly (Supplemental Note; Figure 6.2), but their facial dysmorphism does not overlap sufficiently to make this condition recognizable via their gestalt. Both individuals with macrocephaly displayed prominent broad foreheads, open mouth appearance, and scooped out appearance on lateral view (Figure 6.2). Collectively, the clinical data suggested a remarkable phenotypic variability in our cohort.

We mapped each of the seven identified mutations onto the available crystal structure of human RAC1 (PDB: 3TH5) (Figure 6.1D) to gain insights into their effect on the protein structure and function. The amino acids affected by the seven mutations could be categorized into three groups. (1) Cys18 and Cys157 are located in and adjacent to, respectively, the guanine nucleotide binding site that binds GTP/GDP. These two cysteinethiols can be oxidized by glutathionylation, a post-translational modification that alters GTP binding and exchange activities of RAC1.¹⁷⁷ The p.Cys18Tyr and p.Cys157Tyr substitutions introduce a bulky aromatic residue in place of the thiol group and are expected to impact GTPase activity either by directly interfering with GTP binding or indirectly by abolishing the post-translational modification. (2) Asn39 is part of the switch I motif, while Tyr64 and Pro73 are within and adjacent to, respectively, the switch II motif. Both switch motifs are highly conserved regions involved in the interactions with various GEFs and GAPs (e.g., Rex1, DOCK),¹⁷⁸ mediating the conformational changes for guanine nucleotide exchange (Supplemental Figure 1), and downstream effectors. Missense mutations affecting these residues are likely to impact the protein-protein interactions directly. (3) Val51 has no obvious involvement in GTP binding or interactions with GEF/GAPs. Free energy calculations by FoldX (in the SNPeffect 4.0 server)¹⁷⁹ revealed a reduction of protein stability for these two substitutions ($\Delta\Delta G$ of 2.13 and 1.94 kcal/mol for p.Val51Met and p.Val51Leu, respectively).

RAC1 is known to regulate the spreading of fibroblasts plated onto fibronectin.¹⁸⁰ The genetic data and in silico modeling suggested that the phenotypes are unlikely to result from haploinsufficiency. We reasoned that the *RAC1* mutations identified in this study could be dominant negative or acting as constitutively active. If so, these mutations

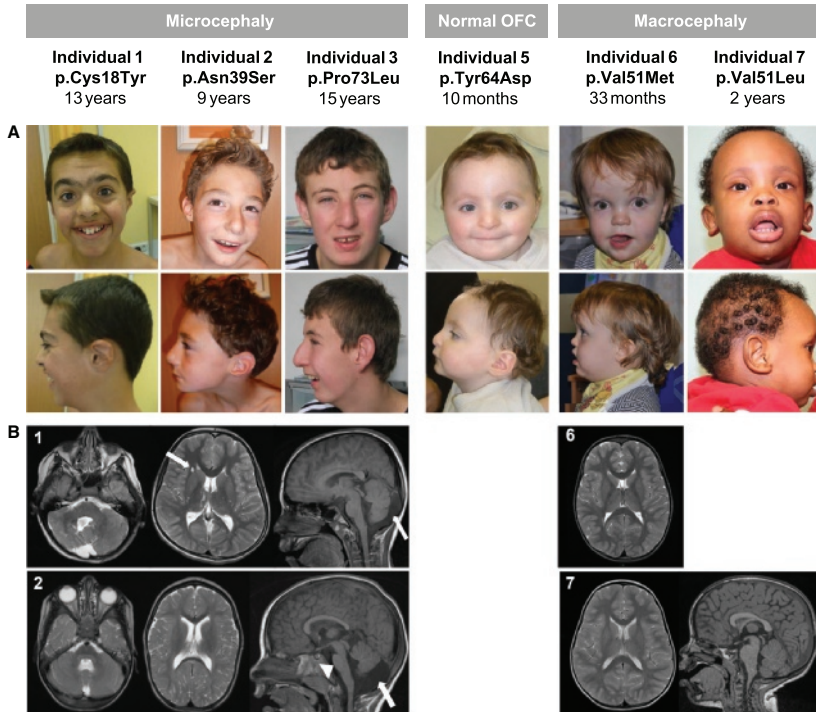


Figure 6.2: Facial dysmorphism and brain abnormalities observed in individuals with RAC1 mutations.

(A) Frontal and lateral photographs of the faces of individuals 1–3 and 5–7. Individuals 1, 2, and 3 were microcephalic, individual 5 had a normal occipital-frontal circumference, and individuals 6 and 7 had macrocephaly. Variable degrees of arched eyebrows, dysplastic ears, prominent nasal bridge, and overhanging columella were noted. Individuals with macrocephaly have prominent broad foreheads, slightly up-slanted palpebral fissures, open mouths, and “scooped out” appearance on lateral view. (B) Panel 1 shows brain MRI images of individual 1: axial T2 images showing dysgenetic vermis and right cerebellar hemisphere (left), cystic lesions in the right frontal deep white matter (middle, arrow), and sagittal T1 images (right) showing enlarged cisterna magna (arrow). Panel 2 shows brain MRI images of individual 2: axial T2 images showing hypoplastic cerebellar vermis (left), slightly enlarged lateral ventricles (middle), and sagittal T1 images (right) showing hypoplastic corpus callosum, pons (arrow head), hypoplastic lower vermis, enlarged 4th ventricle, and cisterna magna (arrow). Panel 6 shows brain MRI image of individual 6: axial T2 image showing non-specific white matter changes in the frontal and parietal lobes. No hydrocephalus or extra-axial fluid was observed. Panel 7 shows brain MRI of individual 7: T2 weighted MRI images. Left: axial image showing very discrete bilateral patchy high intensity signal changes in the deep white matter of the frontal and parietal lobes; right: mid-sagittal section showing, except for megalencephaly, normal brain structures.

should result in changes in fibroblast spreading. To test this hypothesis, we introduced selected *RAC1* mutations in NIH 3T3 fibroblasts. Mammalian expression plasmids encoding GFP-Rac1, GFP-Rac1-T17N, and GFP-Rac1-Q61L were obtained from Prof. Viki Allan (Manchester). A Quikchange Lightning kit (Agilent Technologies) was used to introduce point mutations in the GFP-Rac1 expression plasmid. We generated plasmids encoding 6/7 variants identified in this study (p.Cys18Tyr, p.Asn39Ser, p.Val51Met, p.Tyr64Asp, p.Pro73Leu, and p.Cys157Tyr) along with p.Thr17Asn that is known to have a dominant-negative effect¹⁸¹ and p.Gln61Leu which is known to result in constitutive protein activation.¹⁸² NIH 3T3 fibroblasts were cultured in a 24-well plate and were transfected with

the Rac expression plasmids 24 hr after plating using Eugene 6 reagent (Promega). 48 hr after transfection, the fibroblasts were trypsinized then replated onto fibronectin-coated coverslips. 30 min after re-plating, the coverslips were rinsed with PBS and fixed with 4% paraformaldehyde (PFA). The fixed cells were then permeabilized with 0.1% Triton X-100 in PBS followed by blocking with 1% BSA in PBS. The cells were stained with a rabbit anti-GFP (Invitrogen) antibody (1:500), followed by an Alexa 488 anti-rabbit secondary antibody (Invitrogen) and Alexa568-phalloidin (Invitrogen) before mounting in Prolong Gold (Invitrogen). Cells were imaged on a Nikon A1R confocal microscope, using a 60× 1.4NA oil objective.

Cell circularity values were obtained using ImageJ software. Cell perimeters were identified using the ImageJ “threshold” function and then circularity index ($4\pi \times \text{area}/\text{perimeter}^2$) was calculated using the “analyze particles” function. Circularity datasets were statistically analyzed using one-way ANOVA with Dunnett’s correction for multiple comparisons. At least 50 cells were analyzed for each dataset except p.Asn39Ser, p.Tyr64Asp, p.Cys157Tyr (>40 cells), and p.Gln61Leu (25 cells). Data were pooled from three independent experiments and highly expressing cells were excluded. Morphology was also assessed qualitatively by classifying cells according to their predominant actin protrusion type. Three categories were used: (1) >50% of cell perimeter occupied of filopodia, (2) >50% of perimeter occupied by lamellipodia/ruffles, and (3) mixed protrusions, with neither type occupying 50% of the perimeter.

The majority (51%) of cells transfected with wild-type Rac1 exhibited mixed protrusions, with the remaining cells split almost equally between those in which >50% the perimeter was occupied by filopodia or lamellipodia/ruffles (Figure 6.3A). As observed previously, transfection of the known dominant-negative p.Thr17Asn variant resulted in a filopodia-rich cell perimeter and a significantly reduced circularity index (Figures 6.3A–6.3C).¹⁸⁰ By contrast, expression of the known constitutively active p.Gln61Leu variant resulted in virtually all cells exhibiting large lamellipodia or membrane ruffles and a significantly increased circularity index.¹⁸³ Expression of the p.Cys18Tyr and p.Asn39Ser variants (seen in individuals 1 and 2 with microcephaly) resulted in a phenotype reminiscent of dominant-negative Rac1, with an increase in the proportion of cells rich in filopodia and a reduction in cells rich in lamellipodia/ruffles (Figures 6.3A and 6.3C). Consistently, these mutations exhibited a significantly decreased circularity index (Figure 6.3B). By contrast, cells transfected with the p.Tyr64Asp substitution resulted in a phenotype more reminiscent of constitutive active Rac1, with significantly increased circularity index and a greater proportion of cells exhibiting lamellipodia or ruffles (Figures 6.3A–6.3C). Cells expressing p.Val51Met, p.Pro73Leu, and p.Cys157Tyr all showed a tendency toward increased filopodia and reduced lamellipodia but did not result in a significant change in circularity index relative to cells expressing wild-type Rac, suggesting at most a modest impact on Rac function in these assays (Figures

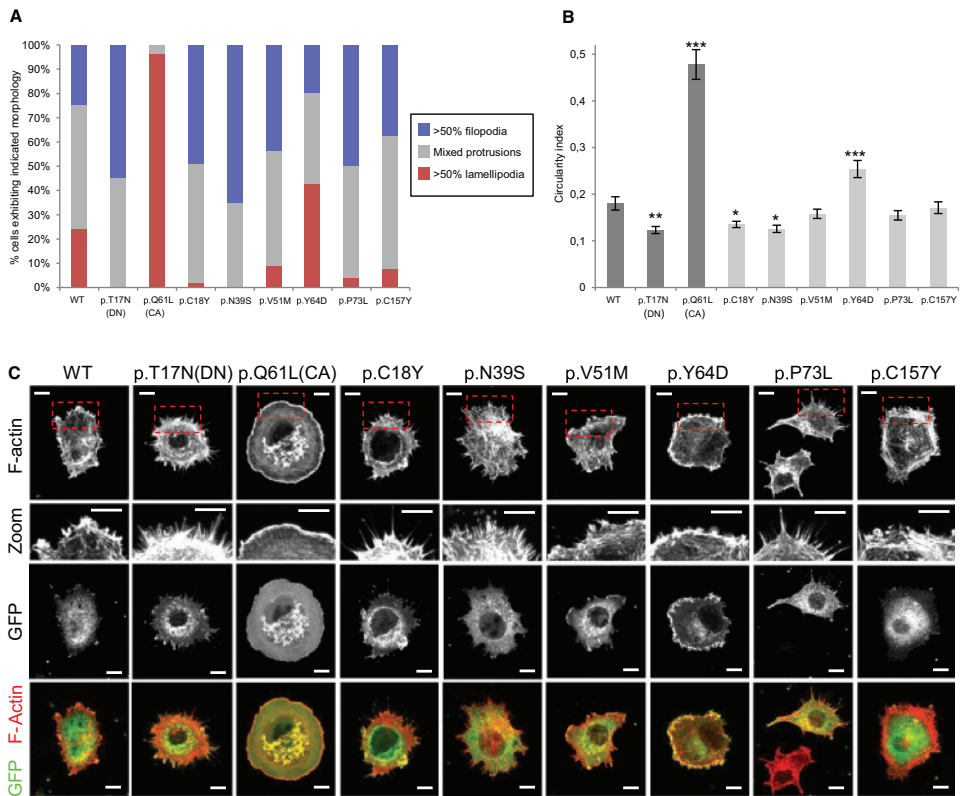


Figure 6.3: Distinct RAC1 mutations genocopy dominant-Negative or constitutively active variants.

(A) NIH 3T3 fibroblasts transfected with 500 ng of DNA per coverslip expressing the indicated Rac1 mutants and were fixed 30 min after plating onto fibronectin and stained with Alexa-568 phalloidin to reveal the actin cytoskeleton and anti-GFP to reveal the expressed construct. Cells were imaged by confocal microscopy and divided into the three indicated categories. At least 50 cells were analyzed for each dataset except p.Asn39Ser, p.Tyr64Asp, p.Cys157Tyr (40 cells), and p.Gln61Leu (25 cells). Data were pooled from three independent experiments. (B) Graph showing mean circularity index of cells expressing indicated Rac1 mutants. Error bars indicate SEM. Datasets statistically analyzed using ANOVA test with Dunnett's correction for multiple comparisons. Asterisks indicate datasets significantly different to WT (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (C) Images of cells expressing the indicated RAC1 mutation. In each case the cell is a representative example of the most common morphological category for that mutation (see A). Top panels show Alexa-568 phalloidin staining to reveal F-actin distribution, while the panels immediately below show a magnified view of the boxed region from the top panels to highlight the morphology of actin protrusions at the cell periphery. The third row of panels show the GFP channel of the same cells to reveal the distribution of the expressed RAC1 mutation, and the bottom panels show a merge of the F-actin (red) and GFP (green) channels. All scale bars indicate 10 μ m.

6.3A–6.3C). All the Rac1 mutant proteins exhibited similar cellular localization to wild-type Rac1 with the exception of p.Tyr64Asp, which appeared to localize strongly to the leading edge of ruffles and lamellipodia (Figure 6.3C).

Next, we sought to explore the *in vivo* effects of RAC1 mutations in zebrafish embryos. Toward this, we identified *rac1a* as the sole zebrafish ortholog with highest degree of homology to the human RAC1 (90% similarity, 90% identity) and used the CRISPR/

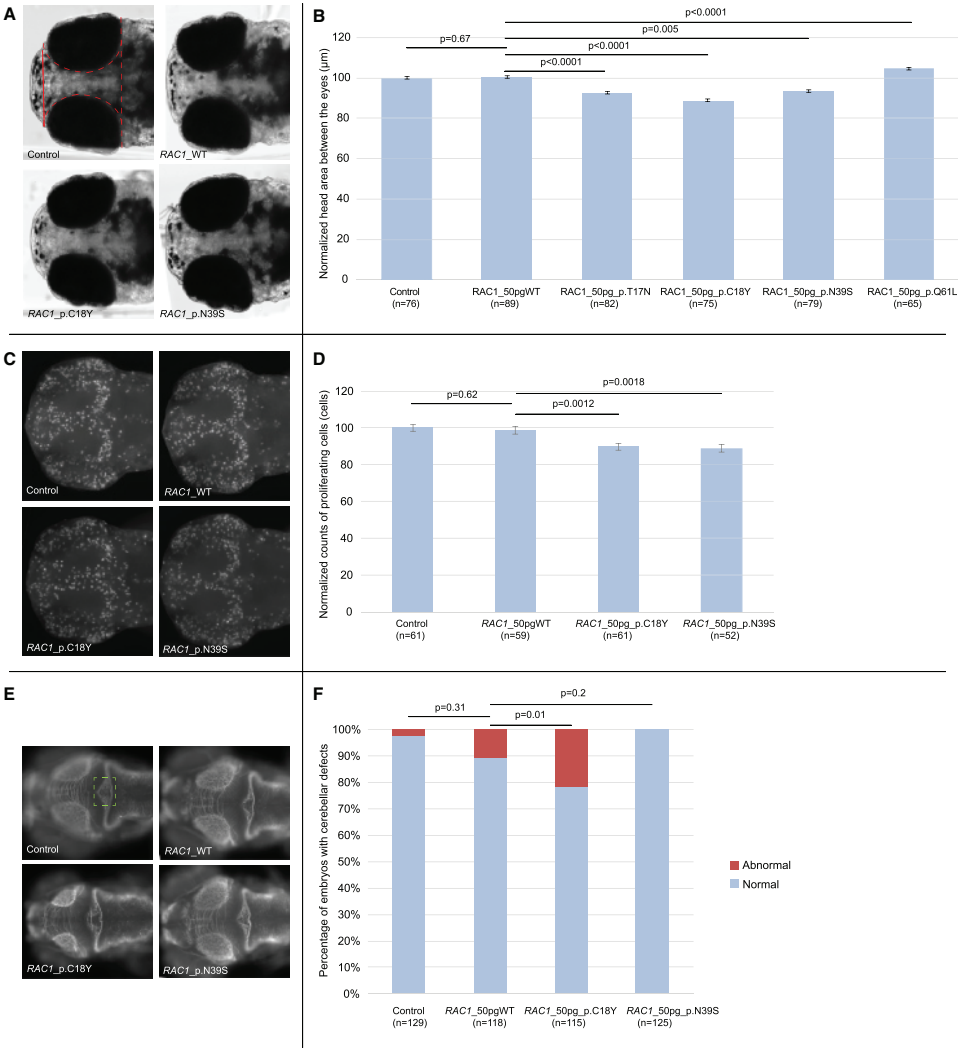


Figure 6.4: Overexpression of human RAC1 p.Cys18Tyr and p.Asn39Ser cause microcephaly, while p.Gln61Leu causes the mirror phenotype in zebrafish embryos.

Functional assessment of de novo variants in RAC1 by in vivo complementation in zebrafish larvae. (A) Dorsal view of 5 days post fertilization (dpf) control and overexpressant larvae. For each experiment, embryos were injected with either WT or mutant RAC1 human mRNA message. The embryos were allowed to grow to 5 dpf and imaged live for head size. The midbrain area between the eyes highlighted with the dashed red line was measured for every imaged embryo, to produce a quantitative score. (B) Bar graph of normalized values showing the quantification of the head size phenotype in control embryos and embryos injected with either WT or mutant human RAC1 message, from two plotted experiments. Statistical analyses were performed by Student's t test. (C) Dorsal view of 2 dpf control and overexpressant larvae, fixed in Dent's fixative and whole-mount immuno-stained with an anti-phospho histone 3 (PH3) antibody that marks proliferating cells. For each experiment, embryos were injected with either WT or mutant RAC1 human mRNA message. The embryos were allowed to grow to 2 dpf and subsequently were fixed, stained, and imaged. Embryos were imaged dorsally and z stacks were acquired every 100 nm. The z stacks were collapsed to produce an extended depth of focus (EDF) image that was then processed for scoring. The PH3-positive cells were counted for every embryo imaged using the ITCN plugin from ImageJ, to produce a quantitative score. (D) Bar graph of normalized values showing the quantification of the proliferating cell count phenotype, across three biological replicas. Statistical analyses for the neuronal proliferation experiments were performed by Student's t test. (E) Dorsal view of 3 dpf

control and overexpressant larvae for *RAC1* WT and mutant conditions. The area of the cerebellum consisting of the neuronal axons that cross the midline is highlighted with a green dashed box. At least 50 embryos per condition were imaged live and evaluated at 3 dpf for the depletion of axons within the area highlighted. (F) Bar graph of cumulative plotted experiments across three biological replicas showing percentages of embryos with cerebellar defects. Statistical analyses for the cerebellar integrity assay using a χ^2 test. Error bars define 95% confidence interval.

Cas9 system to introduce deletions, as a way to explore the phenotypic effects induced by loss of function of *RAC1*. Guide RNAs targeting the *Danio rerio* coding region of *rac1* were generated as previously described.^{184; 185} We observed sequence aberrations in 30% of the evaluated *rac1* clones (Supplemental Figure 2A). Assessment of mosaic F0 embryos injected with a guide against exon 2 did not result in statistically different head size counts between F0 CRISPR and control embryos (Supplemental Figure 2B). No overt morphological changes were observed in the genetically edited embryos.

We next considered a dominant effect of the *RAC1* alleles. To test this hypothesis, we cloned the human wild-type *RAC1* mRNA (GenBank: NM_018890) into the pCS2+ vector and transcribed *in vitro* using the SP6 Message Machine kit (Ambion). The mouse fibroblast spreading assays had indicated that dominant-negative effects could be a common mechanism for *RAC1* mutations. We therefore introduced the two variants (p.Cys18Tyr and p.Asn39Ser) that were shown to be the strongest genocopies of the known dominant-negative mutation using Phusion high-fidelity DNA polymerase (New England Biolabs) and custom-designed primers. Additionally, for sake of comparison, we also introduced the known dominant-negative (p.Thr17Asn) and constitutively active (p.Gln61Leu) mutants. Based on the dose-curve for the titration of the effect that WT *RAC1* had on the head-size phenotype (Supplemental Figure 3), we injected 50 pg of WT or mutant RNA into wild-type zebrafish embryos at the 1- to 4-cell stage. The injected larvae were grown to 5 dpf and imaged live on dorsal view. The area of the head was traced excluding the eyes from the measurements and statistical significance was calculated using Student's t test. All experiments were repeated three times and scored blind to injection cocktail. Injection of RNA encoding p.Thr17Asn, p.Cys18Tyr, or p.Asn39Ser in zebrafish embryos induced a significant decrease in head-size ($p < 0.0005$) compared to controls (Figures 6.4A and 6.4B). Injection of RNA encoding p.Gln61Leu induced a significant increase in head-size ($p < 0.0001$) compared to controls (Figures 6.4A and 6.4B). Importantly, the head-size could not be rescued by co-injection of mutant *RAC1* with WT message (Supplemental Figure 4) arguing further in favor of these *RAC1* mutations acting as dominant alleles *in vivo*, consistent with our results of mouse fibroblast spreading assay.

Driven by the fact that *RAC1* is involved in neuronal proliferation,¹⁷³ we next assessed neuronal proliferation in the brain of embryos injected with WT or a subset of mutant *RAC1* (encoding p.Cys18Tyr or p.Asn39Ser). To do this, the injected embryos were

fixed overnight at 48 hr post fertilization (hpf) in Dent's fixative (80% methanol, 20% DMSO) at 4°C. The embryos were first rehydrated in progressively decreasing methanol solutions, bleached with 10% H₂O₂, 0.5% KOH, and 0.1% Triton-X. After two washes in PBS, the tissue was permeabilized with proteinase K followed by post-fixation with 4% PFA. PFA-fixed embryos were washed first in PBS and subsequently in IF buffer (0.1% Tween-20, 1% BSA in PBS) for 10 min at room temperature. The embryos were incubated in the blocking buffer (10% FBS, 1% BSA in PBS) for 1 hr at room temperature. After two washes in IF buffer for 10 min each, embryos were incubated overnight at 4°C with 1:500 phospho histone 3 (PH3, a marker for proliferating cells) primary antibody (ser10)-R (sc-8656-R, rabbit, Santa Cruz) in blocking solution. After two additional washes in IF buffer for 10 min each, embryos were incubated in the secondary antibody solution (Alexa Fluor goat anti-mouse IgG [A21207, Invitrogen] and donkey anti-rabbit [A21206, Invitrogen], 1:1,000) in blocking solution for 1 hr at room temperature. Proliferating cells were quantified by counting all positive cells on a dorsal view of a 48 hpf embryos, excluding the eyes from the scored area, using the ITCN ImageJ plugin that counts cells with 10 pixel width and 5 pixel minimum distance between them in order to be considered as separate cells. Statistical significance for this assay was established using Student's t test. This assay showed significantly reduced cellular proliferation for both mutants ($p = 0.0012$ for p.Cys18Tyr and $p = 0.0018$ for p.Asn39Ser when compared to WT *RAC1*; Figures 6.4C and 6.4D).

Finally, armed with prior knowledge that *RAC1* has a known role in cerebellar development¹⁸⁶ together with the observation that cerebellar abnormalities were reported in 3/7 individuals in our cohort, we sought to study the effect of the p.Cys18Tyr and p.Asn39Ser variants on cerebellar development of our zebrafish model. For the assessment of the cerebellum, embryos were fixed at 72 hpf in Dent's fixative and subsequently whole-mount stained following the same staining protocol as for neuronal proliferation and using a primary antibody against acetylated tubulin (1:1,000; Thr7451, mouse, Sigma-Aldrich). The embryos were then scored qualitatively assaying the integrity of the cerebellum by scoring for the presence and organization of axons along the midline of the structure (highlighted with a green dashed box; Figure 6.4E) and statistical significance was determined using a χ^2 test. Structural defects in the integrity of the cerebellum consisted of depletion of the axons that cross the midline were observed only upon overexpression of *RAC1* p.Cys18Tyr ($p = 0.01$) but not *RAC1* p.Asn39Ser ($p = 0.2$) (Figures 6.4E and 6.4F). The latter is likely consistent with the more pronounced cerebellar defect observed in individual 1 (who harbors the p.Cys18Tyr variant) (Figure 6.2B).

In summary, we report seven individuals with *de novo* missense *RAC1* mutations and variable developmental delay with additional features. Remarkably, the OCFs observed in this study ranged from -5 SD to $+4.5$ SD. Previously, some chromosomal regions have been associated with both microcephaly (deletions) and macrocephaly (duplications).¹⁸⁷⁻¹⁸⁹

However, it is extremely rare that point mutations within the same gene can cause differences of such magnitude (~10 SD) in head size of affected individuals. It is interesting to note that *RAC1* is involved in mTOR signaling and other disorders in this pathway also result in significant alterations in head size.^{169; 190; 191} The variability of *RAC1* phenotypes appear to be dependent on specific mutations, although the contribution of genetic background cannot be ruled out. *In silico* modeling, mouse fibroblasts spreading assays, and zebrafish experiments demonstrate that some *RAC1* mutations (those encoding p.Cys18Tyr and p.Asn39Ser) genocopy a known dominant-negative mutant (p.Thr17Asn) and result in reduced neuronal proliferation, microcephaly, and cerebellar abnormalities *in vivo*. On the other hand, the *in vitro* effects of p.Tyr64Asp, seen in affected individual 5 with OFC within the normal range, are similar to the known constitutively active *RAC1* mutation. Of note, *in vitro* expression of dominant-negative and constitutively active *Rac1* have been previously shown to cause opposite effects on dendritic growth and morphology.¹⁷⁰ However, the link between the human mutations and the resultant phenotypes is likely to be complex depending on the balance between interactors, regulators, and effectors of *RAC1* signaling.¹⁹² This is likely to be especially true for other *RAC1* mutations identified in this study that could not be clearly classed as being either dominant negative or constitutively active (those encoding p.Val51Met, p.Pro73Leu, and p.Cys157Tyr). Further studies will be required in the future to uncover the precise underlying mechanisms of phenotypic variability of *RAC1* mutations.

Our findings show that mutations in genes encoding members of the RhoGTPases-family can cause DDs. Of note, mutations in *TRIO* (MIM: 601893), a gene that encodes a *RAC1* GEF, have been shown to result in mild intellectual disability.^{193; 194} Interestingly, missense mutations in *RAC*-GEF domain of *TRIO* result in a more severe phenotype with global developmental delay, microcephaly, and reduced *RAC1* activity.¹⁹³ Furthermore, bi-allelic mutations in *HACE1* (MIM: 610876), a known interactor of *RAC1*, have been recently shown to result in an autosomal-recessive syndrome with macrocephaly.¹⁹⁵ Notably, overexpression of a *WAVE* mutant has been demonstrated to partially rescue axon growth in *Rac1* knock-out neurons,¹⁷¹ suggesting that some of the conditions in this group could be potentially treatable. Overall, a potentially treatable sub-category of rare DDs appears to be emerging with *RAC1* as the central player.

Finally, our results show that ultra-rare disorders caused by private, non-recurrent missense mutations, resulting in varying phenotypic effects severity and degrees of severity, are challenging to dissect but can be delineated through focused international collaborations.



7

Variation in a range of mTOR-related genes associates with intracranial volume and intellectual disability

This chapter has been published as:

M.R.F. Reijnders*, M. Kousi*, G.M. van Woerden*, M. Klein, J. Bralten, G.M.S. Mancini, T. van Essen, M. Proietti-Onori, E.E.J. Smeets, M. van Gastel, A.P.A. Stegmann, S.J.C. Stevens, S.H. Lelieveld, C. Gilissen, R. Pfundt, Perciliz L. Tan, T. Kleefstra, B. Franke, Y. Elgersma*, N. Katsanis*, H.G. Brunner*

Nature Communications (2017) 8, 1052

* These authors contributed equally

Abstract

De novo mutations in specific mTOR pathway genes cause brain overgrowth in the context of intellectual disability (ID). By analyzing 101 mTOR-related genes in a large ID patient cohort and two independent population cohorts, we show that these genes modulate brain growth in health and disease. We report the mTOR activator gene *RHEB* as an ID gene that is associated with megalencephaly when mutated. Functional testing of mutant *RHEB* in vertebrate animal models indicates pathway hyperactivation with a concomitant increase in cell and head size, aberrant neuronal migration, and induction of seizures, concordant with the human phenotype. This study reveals that tight control of brain volume is exerted through a large community of mTOR-related genes. Human brain volume can be altered, by either rare disruptive events causing hyperactivation of the pathway, or through the collective effects of common alleles.

Introduction

Many aspects of brain homeostasis, among which are measures of total brain volume, are highly heritable.¹⁹⁶ Genome-wide association studies (GWAS) of brain volume have shown a polygenic architecture in the general population, with individual common genetic variants explaining <1% of phenotypic variance.¹⁹⁷ Neurodevelopmental disorders, such as intellectual disability (ID) and autism spectrum disorder (ASD), have been associated with significant brain overgrowth. In ID, up to 6% of the patients are macrocephalic.²⁸ One of the key regulators of normal brain development is the evolutionarily conserved Ser/Thr protein kinase Mammalian Target Of Rapamycin (*mTOR*). The role of the mTOR pathway in brain development and function has been intensively studied both in vitro and in vivo using different mouse models. In these models, mutations in either the downstream effectors of mTOR, or the most important upstream regulators of mTOR, such as Ras homolog enriched in brain (Rheb), tuberous sclerosis 1 (Tsc1), and Tsc2, have been tested.^{198; 199} Collectively, all studies provide strong evidence that proper mTOR signaling is involved in key aspects of brain development, such as neuronal progenitor maintenance and differentiation (including regulation of neuronal polarity, soma size and neurite outgrowth) and neuronal migration.²⁰⁰⁻²¹¹ In the mature brain, mTOR is an important regulator of synapse formation and synaptic function,²¹²⁻²¹⁵ in particular through its role in regulating protein translation and elongation.²¹⁶⁻²¹⁹ Not surprisingly, hyperactivity of the mTOR pathway in mice can lead to a myriad of phenotypes such as macrocephaly, seizures, and behavioral abnormalities.^{200; 220-225} In contrast, sustained downregulation of the mTOR pathway appears to have little effect on neuronal function and behavior.²²⁶ Findings that the epilepsy and behavioral deficits in mice can be rescued by mTOR inhibitors, offers a broad therapeutic window in which patients can potentially be treated. Indeed, recent studies indicated that mTOR inhibition is a promising treatment for epilepsy in tuberous sclerosis complex (TSC) patients.²²⁷⁻²³¹

Given the large body of evidence implying mTOR function in key aspects of brain development, it is not surprising that hyperactivating, somatic, and germline mutations in components of the PI3K-AKT3-mTOR pathway have been linked with rare ID syndromes associated with (hemi)megalencephaly, focal cortical dysplasia, and epilepsy.^{190; 232-234} We were struck by the apparent recurrence of mTOR-related mutations in ID, the persistent co-morbid megalencephaly and the absence of studies investigating the overall contribution of the mTOR pathway to ID and brain growth. Considering this knowledge gap, we sought to identify deleterious germline mutations in mTOR-related genes, and assess their contribution to the development of ID and megalencephaly. Next, assuming that our findings are not only relevant to rare diseases such as ID, we hypothesized that the pathology of syndromic ID patients represents the extreme end of a more continuous contribution of the mTOR pathway to human brain development and neuroanatomical variance in the population. Our data indeed indicate that mTOR variation significantly contributes to megalencephaly in a large ID cohort and brain size in the population. Furthermore, we present that de novo mutations in a

key regulator of mTOR, *RHEB*, causes severe ID, epilepsy and megalencephaly in humans. By functionally testing the *RHEB* mutations in vertebrate animal models, we show that the specific mutations cause hyperactivation of mTOR, with a concomitant increase in cell and head size, aberrant neuronal migration and induction of seizures, concordant with the human phenotype. The extent of mTOR activation likely affects brain volume in humans. In extreme cases, highly deleterious mutations can lead to profound pathology. For such patients, functional restoration of the pathway through treatment with selective mTOR inhibitors might be of direct clinical utility.

Results

mTOR-related mutations are associated with macrocephaly

To assess the overall burden of mTOR defects to ID, we performed whole-exome sequencing (WES) in a cohort of 826 patients with ID cataloguing de novo mutations (Supplemental Data 7.1) in a set of 101 mTOR-related genes (Supplemental Data 7.2, Figure 7.1). We identified 17 de novo mutations affecting 10 different mTOR-related genes, providing a possible genetic diagnosis in 2.1% of our cohort. Five of the identified genes were known ID genes (*PIK3R1*, *PIK3R2*, *RAF1*, *PPP2R5D*, *MTOR*) and five (*RHEB*, *RAC1*, *PPP2R5E*, *PPP2CA*, *ERK1*) were not associated with ID previously (Figure 7.1, Table 7.1, Supplemental Data 7.3). Three of the five novel genes (*RHEB*, *RAC1*, and *PPP2CA*) showed a significant enrichment for de novo mutations in our patient cohort (Table 7.1, Supplemental Table 7.1). Combining the gene-specific mutation rates of all individual mTOR-related genes, we found a significant enrichment for de novo mutations in mTOR-related genes ($p=3.50e-04$) (Supplemental Table 7.1). Additionally, we found significant spatial clustering of de novo missense variants for a single gene (*PPP2R5D*: $p<1e-07$; permutation test) and a general pattern of spatial clustering across the five genes with recurrent de novo missense variants ($p=0.0057$, Fisher's combined probability test; Supplemental Table 7.2).

To investigate the contribution of mTOR-related mutations on brain overgrowth, we performed a literature analysis of the 101 mTOR genes. This search showed that 23 genes had been previously reported to cause syndromic ID, with the majority (18/23; 78%) being associated with varying degrees of macrocephaly or relative macrocephaly (Supplemental Table 7.3). Motivated by this observation, we collected occipital frontal circumference (OFC) data from 732/826 patients (Supplemental Data 7.4). Macrocephaly was present in 6% of patients in our cohort (47/732 ID patients), a rate comparable to previous reports from an independent cohort³. De novo mutations were identified in 76% of our cohort (553/732 patients; Table 7.2). Among the 35 patients presenting with ID, macrocephaly, and a de novo mutation, we found a significant enrichment ($p=9.084e-09$) for de novo mutations within genes of the mTOR pathway (9/14) compared to genes that operate in mTOR independent pathways (26/539) (Table 7.2). In contrast, microcephaly was not enriched among patients with de novo mutations in mTOR-related genes ($p=0.4228$).

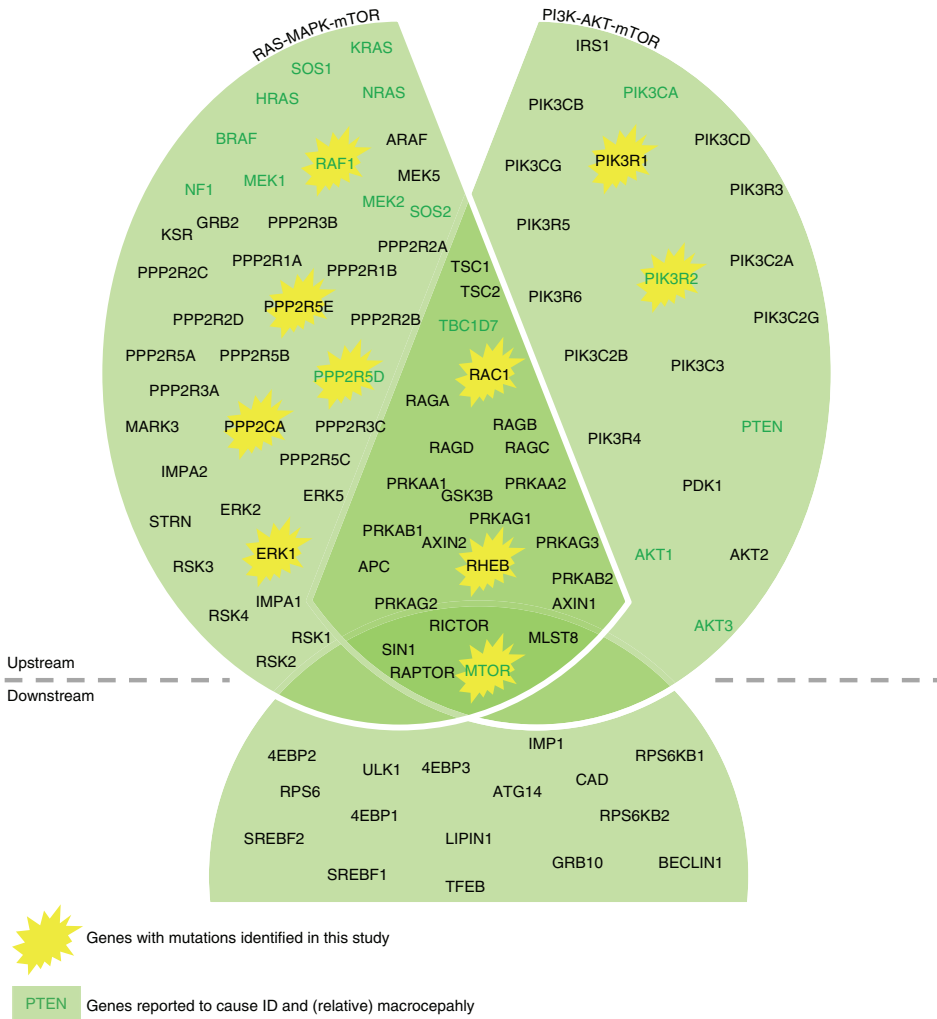


Figure 7.1: Schematic overview of selected mTOR-related genes.

Schematic representation of the genes ($n=101$) included in our mTOR-related gene-set based on three different authoritative publication.²⁴⁹⁻²⁵¹ Both proteins acting upstream of mTOR and proteins acting downstream of mTOR were included. Additionally, we subdivided the total set in two subsets: the RAS-MAPK-mTOR subset and the PI3K-AKT-mTOR subset. In both subsets, downstream genes are included as well. Genes in which we identified de novo mutations in this study were marked with a yellow star, and genes previously reported to cause ID and (relative) macrocephaly are shown in bold and dark green.

mTOR pathway contributes to intracranial volume

Driven by the high frequency of brain overgrowth described in the literature and the strong enrichment of macrocephaly in patients with mutations in mTOR-related genes in our cohort, we tested our set of 101 mTOR pathway genes for an association with intracranial volume (ICV) in the general population (Figure 7.1, Supplemental Table 7.2).

Table 7.1 Identified mutations with bonferroni corrected p-value and occipital frontal circumference (OFC) of patients with de novo mutations in mTOR-related genes

Patient key	Gene	c.DNA	Protein change	Known ID gene	OFC	Bonferroni corrected p-value
1	RHEB	c.202T>C	p.(S68P)	No	> +2.5 SD	4.514e-03*
2	RHEB	c.110C>T	p.(P37L)	No	> +2.5 SD	
3	RHEB	c.110C>T	p.(P37L)	No	> +2.5 SD	
4	RAC1	c.53G>A	p.(C18Y)	No	< -2.5 SD	7.657e-03
5	RAC1	c.116A>G	p.(N39S)	No	< -2.5 SD	
6	PPP2R5E	c.605T>G	p.(V202G)	No	No data	1
7	PPP2CA	c.882dup	p.(R295*)	No	No data	1.696e-02
8	PPP2CA	c.572A>G	p.(H191R)	No	> +2.5 SD	
9	ERK1	c.569T>C	p.(I190T)	No	< -2.5 SD	1
10	PIK3R1	c.1359C>G	p.(N453K)	Yes	Normal	7.662e-02
11	PIK3R1	c.1692C>G	p.(N564K)	Yes	> +2.5 SD	
12	PIK3R2	c.1117G>A	p.(G373R)	Yes	> +2.5 SD	1
13	RAF1	c.1082G>C	p.(G361A)	Yes	Normal	1
14	PPP2R5D	c.1258G>A	p.(E420K)	Yes	No data	7.832e-04
15	PPP2R5D	c.598G>A	p.(E200K)	Yes	> +2.5 SD	
16	PPP2R5D	c.592G>A	p.(E198K)	Yes	> +2.5 SD	
17	MTOR	c.4555G>A	p.(A1519T)	Yes	> +2.5 SD	1

Abbreviations: OFC = occipital frontal circumference; SD = standard deviation

*In the statistical enrichment analysis, the RHEB p.(P37L) variant was considered as a single event.

The final data set contained 76,746 SNPs in 96 autosomal genes (data were unavailable for X-chromosomal *ARAF*, *RPS6KA3*, *RPS6KA6*, *RRAGB*, and *PPP2R3B*). Using the ENIGMA2 data set ($n=13,171$) we found a significant association of the entire mTOR gene set with ICV for the self-contained test ($p_{\text{self-contained}}=0.0029088$) and a suggestive association for the competitive test ($p_{\text{competitive}}=0.054742$). Data from the CHARGE consortium ($n=12,803$) similarly revealed a significant association of the mTOR gene set with ICV for the self-contained test, but not for the competitive test ($p_{\text{self-contained}}=0.00076589$ and $p_{\text{competitive}}=0.22105$, respectively). Meta-analysis of the two data sets, confirmed the significant association of the mTOR gene set with ICV both for self-contained and competitive tests ($p_{\text{self-contained}}=1.3895e-05$, $p_{\text{competitive}}=0.025764$). Post hoc testing of the two major branches of the mTOR pathway separately (RAS-MAPK-mTOR, 76 genes; PI3K-AKT-mTOR, 60 genes; Figure 7.1, Supplemental

Table 7.2 Number of patients with macrocephaly, normal OFC and microcephaly

	Macrocephaly	Normal OFC	Microcephaly
Patients with clinical data (n=732)	47	580	105
Patients with de novo mutation(s) (n=553)	35	442	76
Patients with de novo mutation in mTOR-related gene (n=14)	9	2	3
Patients with de novo mutation in gene not related to MTOR (n=539)	26	440	73

Abbreviation: OFC = occipital frontal circumference

Table 7.4) showed stronger association of PI3K-AKT-mTOR than RAS-MAPK-mTOR with ICV (PI3K-AKT-mTOR: $p_{\text{self-contained}}=0.00092471$, $p_{\text{competitive}}=0.0079133$; RAS-MAPK-mTOR: $p_{\text{self-contained}}=2.2885\text{e-}07$, $p_{\text{competitive}}=0.068983$). The role of the PI3K-AKT-mTOR pathway in volumetric variation of the brain was further strengthened by testing the previously described Reactome_PI3K_AKT_activation gene-set²³⁵ (35 genes; $p_{\text{self-contained}}=3.8649\text{e-}13$; $p_{\text{competitive}}=0.00028957$; Supplemental Table 7.5). Not all 96 genes in the gene-set analysis showed significant association with ICV individually. The most strongly associated individual gene was *AKT3* ($P=2.22\text{E-}05$) and in total, 18 genes of the mTOR gene set including *APC* ($P=0.00042$), and the new ID gene *RHEB* ($P=0.0041$), showed nominally significant association with ICV (Supplemental Figure 7.1; Supplemental Table 7.6).

RHEB mutations cause increased neuronal cell and head size

Two of the three individuals with de novo *RHEB* mutations were siblings and carried the same heterozygous p.(Pro37Leu) mutation, while a sporadic individual carried the p.(Ser68Pro) allele. The p.(Pro37Leu) mutation was not identified in either parent, suggesting parental gonadal mosaicism (Figure 7.2A). The *RHEB* mutations are located in the RAS domain (Figure 7.2B and C) and are absent from ExAC, EVS, or our internal clinical exome databases. All three individuals (Figure 7.2D) with de novo *RHEB* mutations had short stature (-2 to -3SD) and early brain overgrowth with pronounced macrocephaly during childhood ($+2.5/+3\text{SD}$). They had severe to profound ID with hypotonia, as well as autism spectrum disorder. Two of three individuals were reported to have epilepsy. No epileptic episodes were noted for the third patient, but EEG recordings were suggestive of epileptic discharges (Supplementary Note, Supplemental Table 7.7). Brain magnetic resonance imaging (MRI) evaluation of the patient with the p.(Ser68Pro) allele, confirmed megalencephaly with broad frontal lobes and mild dilatation of the lateral ventricles. The MRI scan further showed a thickened rostrum of the corpus callosum and small splenium, and mild hypoplasia of the lower cerebellar vermis (Figure 7.2E).

We selected the *RHEB* mutations to obtain experimental evidence for our hypothesis that de novo changes in mTOR-related genes are likely due to a gain-of-function

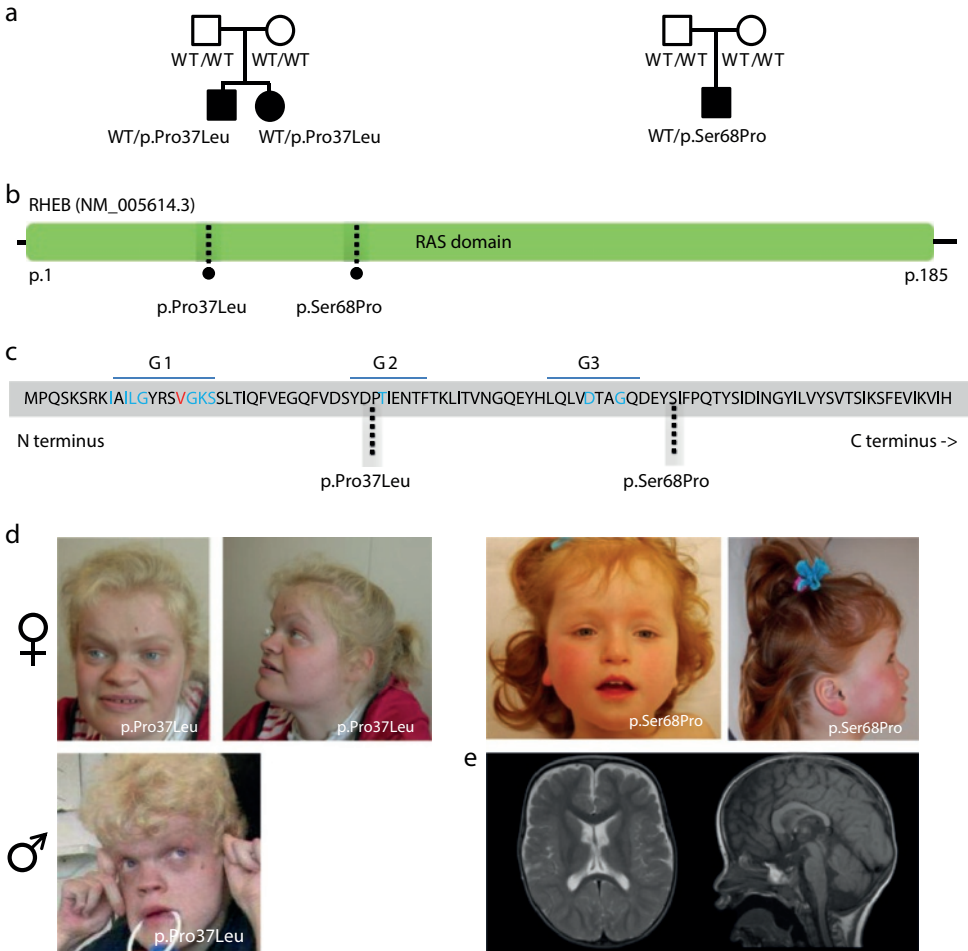


Figure 7.2: De novo mutations in RHEB cause an ID syndrome associated with megalencephaly. (A) Pedigree structure, disease status and genotype information for patients with changes in RHEB. (B) Schematic of the RHEB protein and the RAS domain. (C) Schematic of the N-terminal portion of the human RHEB protein. G-box residues characteristic of RAS superfamily proteins' are shown in blue; highly conserved residue conserved in 90% of the RAS superfamily members are shown in red. Dotted lines are showing the amino-acid residues mutated in patients described in the context of this study. (D) Photographs of the probands carrying de novo variants in RHEB. (E) MRI images (left: axial, T2-weighted; right: sagittal midline, T1-weighted) of the proband (age 1 year, 9 months) carrying the de novo RHEBp.S68P variant, showing macrocephaly, megalencephaly, broad frontal lobes, mild dilatation of lateral ventricles, large rostrum of corpus callosum and mild hypoplasia of the lower cerebellar vermis. No cortical malformations have been observed.

mechanism, resulting in hyperactivation of mTOR, as previously shown for other syndromic neurodevelopmental cases associated with macrocephaly. We first tested in vitro whether the RHEB de novo changes have an impact on overall mTOR activity levels. Given that mTORC1 regulates cell size,^{205; 236} we used primary hippocampal neuron soma size as a readout to assess differences between RHEB-WT overexpressing vs. RHEBp.P37L and

RHEBp.S68P overexpressing neurons. A significant increase in soma size was detected already in *RHEB*-WT transfected neurons, suggesting that *RHEB* is a highly dosage sensitive gene, likely causing hyperactivation of the mTOR pathway.²³⁷ Overexpression of the RHEB mutant proteins caused an increase in soma size, confirming that these mutations do not cause a loss of function (Figures 7.3A and B, one-way ANOVA, $p < 0.0001$, $F(3,260) = 50.35$; control vector vs. RHEB-WT: $p < 0.0001$; control vector vs. RHEBp.P37L: $p < 0.0001$; control vector vs. RHEBp.S68P: $p < 0.0001$ by Tukey's multiple comparisons test). Notably, overexpression of RHEBp.P37L had the strongest effect inducing a significantly pronounced increase in soma size compared to RHEB-WT ($p < 0.05$) and RHEBp.S68P ($p < 0.05$).

We next sought to evaluate the relevance of these variants in the development of neuroanatomical phenotypes in a developing zebrafish *in vivo* model. Toward this, we identified the sole zebrafish *rheb* ortholog (96% similarity, 91% identity). First, we corroborated that the variants identified are not acting through a loss of function mechanism by generating a CRISPR-Cas9 system to introduce deletions. Assessment of head size in mosaic F0 embryos injected with a guideRNA against exon 3 showed microcephaly in two biological replicates, which was opposite to the phenotype observed in the patients (Supplemental Figure 7.2). We next evaluated the effect of the *rheb* alleles on head size under a gain of function and mTOR hyperactivating paradigm, as suggested through our *in vitro* studies. To test this hypothesis, we injected human WT or mutant *RHEB* mRNA into 1- to 4-cell stage zebrafish embryos. Expression of WT human *RHEB* induced a significant increase in the headsize area of 5 dpf larvae ($p = 0.0013$). Overexpression of either RHEBp.P37L or p.S68P, also resulted in significantly increased headsize, reminiscent of the megalencephaly seen in our patients ($p < 0.0001$ for either mutant allele when compared to WT *RHEB*; Figures 7.3D and E). This finding was reproducible across three independent biological replicates.

Rapamycin rescues neuronal soma and head size defects

Antagonists of the mTOR pathway, such as rapamycin, can ameliorate some of neurological deficits associated with mTOR hyperactivity²²⁷⁻²³¹. To evaluate whether this is true for the *RHEB* activating mutations described here, we treated the neurons with 20nM rapamycin or vehicle 1 day after transfection for 3 days and assessed neuronal soma size. We found that while the soma size of RHEB-WT overexpressing neurons nominally decreased, a statistically significant reduction of neuronal soma size was observed for both RHEBp.P37L and RHEBp.S68P and a trend in the same direction was seen for RHEB-WT (Figure 7.3C, two-way ANOVA, effect of treatment $p < 0.0001$, $F(1,89) = 16.29$; RHEBp.P37L vehicle vs. RHEBp.P37L Rapamycin $p < 0.01$; RHEBp.S68P vehicle vs. RHEBp.S68P Rapamycin: $p < 0.05$; RHEB-WT vehicle vs. RHEB-WT Rapamycin: $p = 0.1$; by Bonferroni's multiple comparisons test). Taken together, these data show that overexpression of either wild-type or mutant *RHEB* induces an mTOR-dependent increase in soma size and that

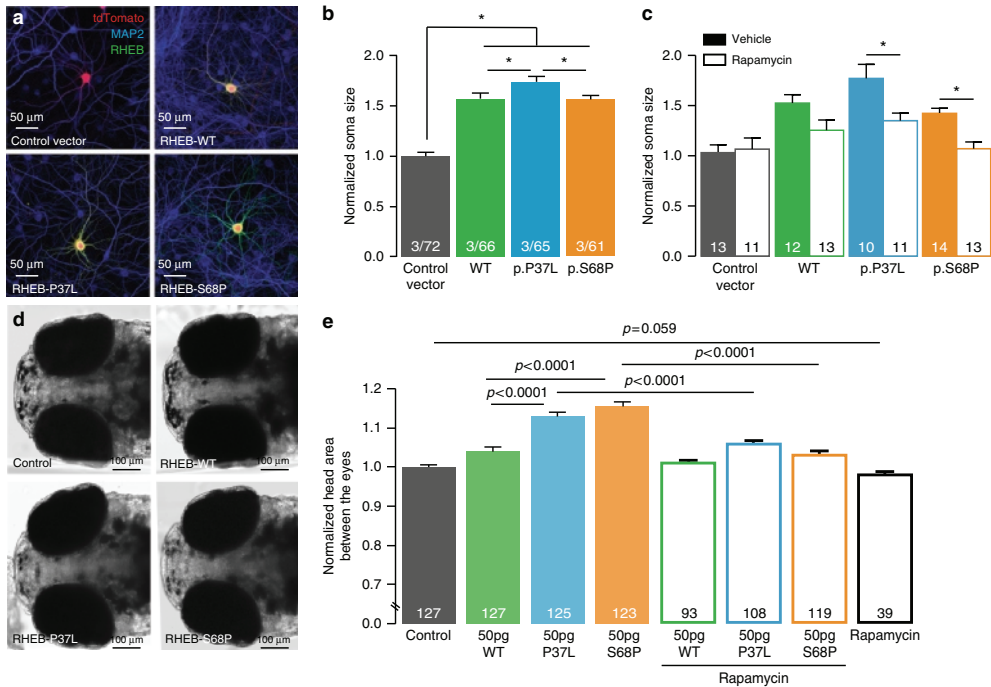


Figure 7.3: De novo mutations in RHEB increase soma size and headsize, phenotypes that can be rescued by rapamycin. (A) Representative confocal images of hippocampal neurons transfected with control vector, RHEB-WT, RHEBp.P37L or RHEBp.S68P. (B) Summary bar graph of soma size measured for each condition and normalized to the control vector. (C) Summary bar graph of soma size measured with and without rapamycin. Data are presented as mean±SEM. Numbers depicted in the bar graph in b represent the number of independent cultures/total number of cells analyzed and in c number of cells analysed. Statistical significance was assessed by one-way ANOVA followed by Tukey's post hoc test (* $p<0.01$). (D) Dorsal view of 5 dpf control and overexpressant zebrafish larvae. For each experiment, embryos were injected with either WT or mutant RHEB human mRNA message. The area between the eyes was measured for every embryo, to produce a quantitative score. (E) Bar graph showing the quantification of the headsize phenotype in control embryos and embryos injected with either WT or mutant human RHEB with and without rapamycin. The graph represents cumulative plotted experiments across three biological replicas. Statistical analyses were performed by Student's t-test.

this phenotype can be rescued through the administration of the known mTOR antagonist rapamycin. Grounded on the *in vitro* observations, we next explored the possibility of rapamycin serving as a putative therapeutic agent *in vivo*. Toward this, we co-injected rapamycin together with WT or mutant RHEB mRNA in zebrafish embryos and we compared the embryos' head size at 5 dpf. Rapamycin alone did not induce any appreciable neuroanatomical pathologies, or indeed any other overt morphological phenotype(s) (Figure 7.3E). In contrast, rapamycin sufficiently and reproducibly rescued the macrocephalic phenotype induced by both RHEBp.P37L and RHEBp.S68P, suggesting that suppression of mTOR hyperactivity might present a therapeutic target for disease amelioration (Figure 7.3E).

RHEB mutations affect neuronal migration and induce seizures

Previous studies have shown that mTOR signaling is not only involved in cell morphology and growth, but also plays a role in neuronal migration. Increased mTOR activity *in vivo*, induced either by overexpression of a constitutively active RHEB or by inactivating mutations in the *Tsc1* or *Tsc2* genes, two negative regulators of *RHEB*, causes neuronal migration defects²⁰⁶⁻²¹¹. We performed in utero electroporation at E14.5 to induce the in vivo overexpression of RHEB-WT, RHEBp.P37L, and RHEBp.S68P, and tested the effect of the *RHEB* mutations on neuronal migration in the still developing somatosensory cortex of P0 pups.^{238; 239} Although in neuronal cultures overexpression of RHEB-WT and RHEB mutants increased soma size equally, the results obtained in vivo showed notable differences between these proteins. We observed that while cells transfected with the control vector efficiently migrated to the cortical plate (CP), cells transfected with RHEB-WT could be found in all the layers of the cortex (Figure 7.4A). Strikingly, the majority of cells transfected with either RHEBp.P37L or RHEBp.S68P remained in the subplate (SP), indicating more severe migration deficits compared to RHEB-WT overexpression (Figure 7.4A). At P7, when the cortical layers are more defined, the difference between RHEB-WT and RHEBp.P37L or RHEBp.S68P was even more striking (Figure 7.4B). Analysis of the number of tdTomato-positive cells present in the different cortical layers showed a significant difference between the four different conditions (two-way repeated measure ANOVA, effect of interaction: $F(27,180)=13.73$, $p<0.0001$), consistent with our previous in vitro and in vivo studies that the mutations in *RHEB* are gain-of-function hyperactivating mutations. Consistent with our findings in primary neuronal cultures, post hoc analysis revealed that the RHEBp.P37L mutation yielded the strongest effects among evaluated conditions (Figure 7.4C).

Neuronal migration deficits are often linked with seizures and ID.²⁴⁰ Additionally, the link between an epileptogenic phenotype and hyperactivity of the mTOR pathway has been established from studies in both human and mice.^{222; 224; 241} Interestingly, knockdown of the *TSC1* gene, a negative regulator of the mTOR pathway, in just a subset of cortical neurons reduces the threshold for seizure induction.²⁰⁶ Careful monitoring of the in utero electroporated mice, revealed that overexpression of RHEB-WT, RHEBp.P37L, and RHEBp.S68P resulted in spontaneous tonic-clonic seizures starting at P20 (Supplemental Movie 7.1). Seizures were particularly common in mice expressing mutant RHEB: whereas 20% (2/10) of mice expressing RHEB1-WT developed epilepsy, all (7/7) mice expressing RHEBp.P37L and 83% (5/6) of mice expressing RHEBp.S68P developed spontaneous seizures (Figure 7.4D). Consistent with our findings from primary neuronal cultures and neuronal migration following in utero electroporation, the RHEBp.P37L allele was shown to have the strongest effect, as the mice expressing this allele also showed a significantly earlier onset of seizures (log-rank (Mantel-Cox) $p<0.01$ compared to RHEBp.S68P; Figure 7.4D). Taken together, our *in vivo* results further corroborate the conclusion that the missense mutations in *RHEB* act as dominant activating mutations.

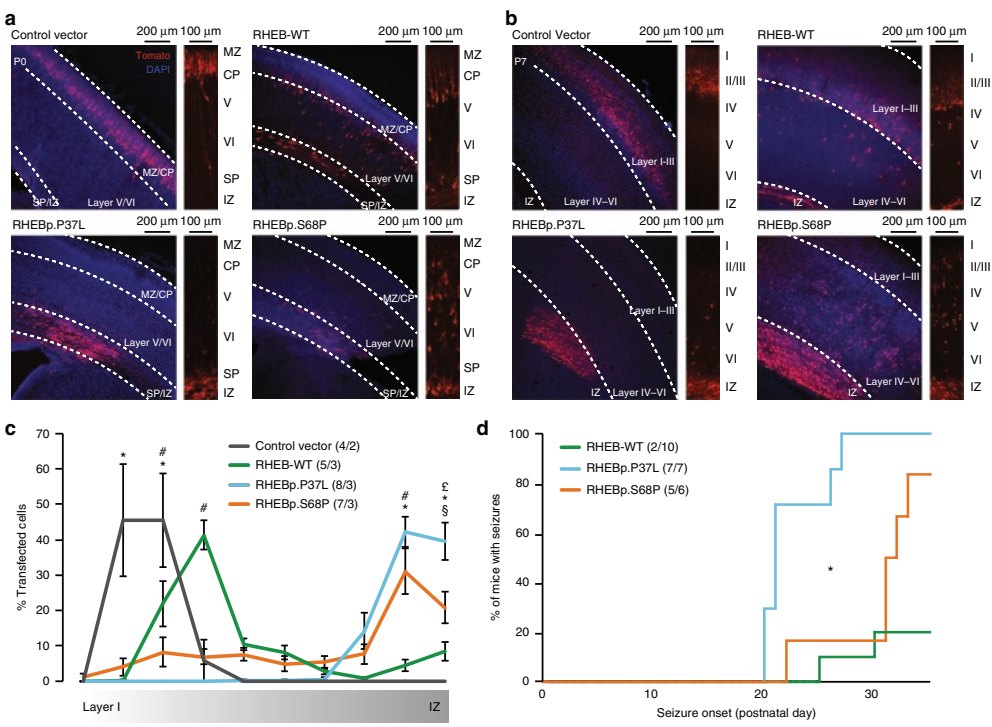


Figure 7.4: Overexpression of RHEB mutants in vivo causes deficits in neuronal migration and seizures in mouse. (A,B) Representative images of E14.5 in utero electroporated P0 brains (a) or P7 brains (b), with an enlargement showing the migratory pathway of the transfected cells (tdTomato+) from the intermediate zone (IZ) and subplate (SP) to the more superficial layers of the cortex (CP=cortical plate and MZ=marginal zone). (C) Quantification of the neuronal migration pattern observed in different conditions. Data are presented as mean±SEM. Statistical significance was assessed by two-way repeated measure ANOVA followed by Bonferroni's post hoc test (for bins 2–4: *indicates significant difference between control vector and the different RHEB conditions ($p<0.0001$); #indicates significant difference between the RHEB-WT and all other conditions ($p<0.0001$); for bins 9 and 10: *indicates significant difference between control vector and RHEBp.P37L and RHEBp.S68P; #indicates significant difference between RHEB-WT and RHEBp.P37L and RHEBp.S68P ($p<0.0001$); §indicates significant difference between RHEB-WT and RHEBp.P37L ($p<0.0001$); ¶indicates significant difference between RHEBp.P37L and RHEBp.S68P ($p<0.001$)). (D) Kaplan–Meier graph representing onset of tonic–clonic seizures in successfully targeted mice. The insert legends of the graph show $N_{pictures}/N_{mice}$ (c) or $N_{seizure}/N_{total}$ (d).

Discussion

Here we studied the contribution of mTOR-related genes to ID and brain overgrowth in 826 ID patients unselected for any other phenotypic features and found 17 germline de novo mutations in genes related to mTOR, providing a possible genetic diagnosis for 2.1% of our cohort. We show that genes encoding components of the mTOR pathway, contribute to rare and common alleles that impact brain volume and provide insight into neurodevelopmental processes mediated through mTOR hyperactivity and outlook to potential treatment options for a subset of patients with ID.

A significant fraction of patients harboring a de novo mutation in mTOR-related genes was

observed to be macrocephalic. The link between mTOR mutations and ID and/or head size differences has already been established through numerous studies that identified genes such as *AKT3*, *PIK3CA*, *PPP2R5D*, and recently *MTOR* itself.^{190; 232; 233; 242} In fact, from the 23 genes previously reported to cause ID among our gene-set of 101 mTOR-related genes, most (18/23, 78%) have been associated in the literature with macrocephaly or relative macrocephaly. Our study significantly extends these findings: of the 35 patients with macrocephaly in the complete cohort, 9 patients (26%) harbored a de novo mutation in mTOR-related genes. As such, genes in this pathway should be carefully evaluated in patients with ID and macrocephaly.

Motivated by the high frequency of brain overgrowth in previous reports of mTOR-related syndromes, and in patients with mutations in mTOR-related genes in our cohort, we reasoned that the highly penetrant activating alleles that we identified de novo might represent only a fraction of alleles associated with severe neurocognitive disorders, and that more common and less penetrant alleles might be associated with head and brain growth in the general population. Indeed, a combined analysis of common variants of all 96 autosomal genes in the mTOR-related gene-set showed significant association with ICV in two large imaging genetics samples from the CHARGE and ENIGMA consortia, confirming our initial hypothesis. Interestingly, the PI3K-AKT-activation pathway (35 genes from the reactome gene-set) was recently shown to be among the most strongly associated pathways for ICV in an enrichment analysis testing 671 Reactome gene-sets using the same cohorts.²³⁵ Our analyses support and expand this conclusion by testing a different, carefully selected gene-set (only 15 out of 96 genes overlapping). Taken together, our data support a model by which mTOR-related genes, including the newly discovered ID gene *RHEB*, contribute to variation in brain growth, through common and rare genetic variants, in health and disease. Our observations corroborate, how rare disorders can inform biological mechanisms underlying common traits in the general population.

There is ongoing debate on the precise genetic composition of gene-sets. Gene-set databases, such as KEGG, Ingenuity, and others, all differ in their coverage of specific biological pathways and their functional annotations. In line with this observation, the number of genes mapped to pathways may also vary greatly across the databases.²⁴³ For mTOR-related gene-sets, inclusion in databases is incomplete, with key proteins and protein complexes such as *RAC1*, *RAG*, *MEK*, and *PP2A* missing. For this reason, we used three authoritative reviews on mTOR signaling describing both upstream and downstream interactors of mTOR and then used additional evidence from the literature to subdivide various protein complexes into their constituent proteins and genes. Therefore, our selection of 101 mTOR-related genes might be incomplete and additional genes are likely to be involved in mTOR signaling. For this reason and because of limitation of our methods to detect reliably somatic mosaicism, a mechanism thought to be a significant contributor to

mutation burden in this pathway,^{190; 232; 233} we postulate that the diagnostic rate within our cohort (2.1%) might represent the lower bound of the estimate.

In this study, we identified de novo mutations in both known ID genes (*PIK3R1*, *PIK3R2*, *RAF1*, *PPP2R5D*, *MTOR*) and novel candidate ID genes (*RHEB*, *RAC1*, *PPP2R5E*, *PPP2CA*, *ERK1*). For the most frequently mutated gene, *RHEB*, we show that hyperactivating mutations cause an ID syndrome with brain overgrowth and epilepsy. The finding that these mutations are hyperactivating, is in line with the observation that loss of RHEB activity does not result in overt neurological phenotypes in *Rheb* mutant mice.²²⁶ Several mechanisms, such as increased proliferation, increased soma size and reduced apoptosis are known to have a role in the development of megalencephaly.²⁴⁴ We observed a significant increase in soma size upon overexpressing WT and mutant RHEB alleles in vitro. Since RHEB is the canonical activator of mTOR, this finding is consistent with other reports that have highlighted mTOR as a main regulator of cell size.^{200; 205; 209; 245} In vivo, we postulate that the increased soma size might represent one of the mechanisms through which macrocephaly occurs, as the zebrafish embryos injected with mutant *RHEB* were phenotypically concordant with the human patients. Further dissecting the pathomechanism(s) underlying *RHEB*-associated ID, we showed severe neuronal migration defects in mouse embryos electroporated with mutant *RHEB* and an increased incidence of epileptogenic activity postnatally. These findings are reminiscent to what has been observed for mutations of *MTOR* itself. Constitutive activation of mTORC1 causes enlarged neuronal somata in rodent neurons, and focal cortical expression of *MTOR* mutations has been reported to disrupt neuronal migration and to cause spontaneous seizures by in utero electroporation in mice.^{233; 234; 246} This observation shows that activating mutations in different genes of the mTORC1 branch of the mTOR pathway act through convergent mechanisms and have similar phenotypic outcomes.

Based on these observations, we reasoned that patients with activating *RHEB* changes might be able to benefit from therapies that result in a reduction of mTOR activity, such as rapamycin. Indeed, we here showed that suppression of mTOR levels through the administration of the mTOR antagonist rapamycin can significantly and reproducibly prevent both the neuronal soma size phenotype in vitro and the macrocephalic phenotype in vivo. Recent studies have reported successful implementation of mTOR inhibitor treatment in individuals with TSC-associated epilepsy and brain tumors.²²⁷⁻²³¹ In a conceptually similar paradigm, fibroblasts from a patient with an mTOR activating *PIK3CA* change, were treated successfully with the PI3K inhibitors wortmannin or LY294002, which abrogated the overactivation of the pathway.²⁴⁷ It is premature to advocate the use of rapamycin in patients with ID and mutations in all mTOR-related genes, not least because of the potentially adverse effects induced by prolonged exposure to this agent.²⁴⁸ However, we speculate that targeted administration of mTOR inhibitors (rapamycin, wortmannin, everolimus as well as currently

emerging second-generation drugs), perhaps during critical postnatal neurodevelopmental windows, might be of significant benefit to patients. In that context, rapid molecular diagnosis in both known ID genes and candidate ID genes, would be a critical component of the treatment decision process.

In conclusion, our data show that a large number of mTOR-related genes together modulate human brain volume in the population. Severe disruption of such mTOR-related genes can cause intellectual disability and brain overgrowth, most likely through mTOR hyperactivation.

Material and methods

Subjects and mutation analysis

We evaluated a cohort of 826 patients with ID, who had undergone diagnostic trio WES at Radboud University Medical Center (Radboudumc). We included 820 simplex patients described previously, as well as three sib pairs excluded from the earlier study.¹²⁶ Diagnostic WES was approved by the medical ethics committee of the Radboud University Medical Center (Commissie Mensgebonden Onderzoek), Nijmegen, The Netherlands (registration number 2011-188). Written informed consent was obtained from all individuals or their legal guardians. We collected all available clinical information and performed deep phenotyping of individuals with a de novo mutation in *RHEB*. Consent for publication of photographs was obtained. Brain images were re-evaluated where available.

Selection of mTOR-related genes

We focused on the two well-described, convergent pathways in which mTOR acts as key regulator: the PI3K-AKT-mTOR pathway and the RAS-MAPK-mTOR pathway. We defined a list of 101 mTOR-related genes based on three authoritative reviews on the mTOR regulators.²⁴⁹⁻²⁵¹ Protein complexes were mapped to single proteins and genes based on information available in the literature. The final list contains 101 mTOR-related genes: 96 map on autosomes and five map on the X-chromosome.

Identification of mutations and collection of OFC data

From our cohort of 826 patients with ID, we selected all de novo mutations that affect mTOR-related genes. All mutations were confirmed by Sanger sequencing. mTOR-related genes were considered to be known ID genes, if present in our recently published list containing over 1500 known ID genes.¹²⁶ We performed a literature search by querying Pubmed to investigate which of the known mTOR-related ID genes have been associated with large or small head size. Within our cohort of 826 patients, individuals were classified as microcephalic (OFC < -2.5SD), macrocephalic (OFC > +2.5SD), normocephalic (OFC between -2.5SD and +2.5SD) or unknown. We used Fisher's Exact test to calculate enrichment of macrocephaly in mTOR-related mutation carriers. The significance threshold was set at $p < 0.05$.

Gene-based enrichment

To assess whether mTOR-related genes were significantly enriched for functional de novo mutations in our cohort, we tested each of the 101 genes using a statistical model as described previously.¹²⁶ For this statistical enrichment analysis, the RHEB p.(P37L) variant was considered as one single event. Multiple testing correction was performed by the Bonferroni procedure based on 101 tested genes. Additionally, we tested whether the mTOR pathway as a whole was enriched for functional de novo mutations in our cohort by combining the gene-specific mutation rates of all individual genes in the pathway.

Clustering analysis

Clustering analysis was performed by generating the full cDNA for the respective RefSeq genes. To increase the statistical power of the spatial clustering of the recurrently mutated genes, we added de novo missense variants from the denovo-db²⁵² annotated by our in-house pipeline (Supplemental Data 7.5). The locations of observed de novo missense mutations were randomly sampled 100,000 times over the cDNA of the gene after which the distances (in base pairs) between the mutations were normalized for the total coding size of the respective gene. The geometric mean (the n th root of the product of n numbers) of all mutation distances between the mutations was taken as a measure of clustering. A pseudocount (adding 1 to all distances and 1 to the gene size) was applied to avoid a mean distance of 0 when there were identical mutations. To assess overall clustering of the set of genes, we used Fisher's combined probability test to combine the 5 p -values of individual genes. To avoid a possible bias introduced by highly significant p -values (e.g., gene *PPP2RD5*), we calculated the combined p -value on deflated p -values where all values smaller than 0.05 were set to 0.05.

ENIGMA and CHARGE study populations and data description

This study reports data on 25,974 subjects of Caucasian ancestry from 46 study sites that are part of the Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA)²⁵³ consortium (13,171 subjects) and Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE; 12,803 subjects).²⁵⁴ Briefly, the ENIGMA consortium brings together numerous studies, mainly with case-control design, which performed neuroimaging in a range of neuropsychiatric or neurodegenerative diseases, as well as healthy normative populations. The CHARGE consortium is a collaboration of predominantly population-based cohort studies that investigate the genetic and molecular underpinnings of age-related complex diseases, including those of the brain. An overview of the demographics and type of contribution for each cohort is provided in Supplemental Table 7.8 (Table adapted from original publication by Adams et al.²³⁵). Written informed consent was obtained from all participants. The study was approved by the institutional review board of the University of Southern California and the local ethics board of Erasmus MC University Medical Center. Procedures of whole-genome genotyping, imputation, MRI,

GWAS, and meta-analysis are summarized in Supplemental Methods (adapted from original publication²³⁵). The meta-analysis data from the recent ENIGMA2 and CHARGE studies of ICV were available as genome-wide summary statistics, including genome-wide single-nucleotide polymorphism (SNP) data with corresponding p -values. The ENIGMA consortium has completed a meta-analysis of site-level GWAS in a discovery sample of 13,171 subjects of European ancestry.^{197; 253} Access to the summary statistics of ENIGMA can be requested via their website (<http://enigma.ini.usc.edu/download-enigma-gwas-results/>). The CHARGE consortium has completed meta-analysis of site-level GWAS in a discovery sample of 12,803 subjects of European ancestry.²³⁵ Genome-wide summary statistics of the CHARGE consortium has been requested by the principal investigator of the study described by Adams et al.²³⁵ For both data sets, only SNPs with an imputation quality score of $RSQ \geq 0.5$ and a minor allele frequency ≥ 0.005 within each site were included. Procedures of whole-genome genotyping, imputation, magnetic resonance imaging (MRI), GWAS, and meta-analysis of the cohorts are summarized in the Supplemental Methods.

Gene-based and gene-set analyses

Gene-based and gene-set analyses were performed using the Multi-marker Analysis of GenoMic Annotation (MAGMA) software package (version 1.02).²⁵⁵ First, gene-based p -values were calculated using a symmetric 100kb flanking region for each cohort separately for the 96 autosomal genes in the mTOR pathway. Genome-wide SNP data from a reference panel (1000 Genomes, v3 phase1)²⁵⁶ was annotated to NCBI Build 37.3 gene locations using a symmetric 100kb flanking window, and both files were downloaded from <http://ctglab.nl/software/magma>. Next, the gene annotation file was used to map the genome-wide SNP data from the different studies (ENIGMA2 and CHARGE), to assign SNPs to genes and to calculate gene-based p -values for each cohort, separately. Since data from the genome-wide association analyses only included autosomal SNPs, five genes located on the X-chromosome were omitted from the analysis. For the gene-based analyses, single SNP p -values within a gene were transformed into a gene-statistic by taking the mean of the χ^2 -statistic among the SNPs in each gene. To account for linkage disequilibrium (LD), the 1000 Genomes Project European sample was used as a reference to estimate the LD between SNPs within (the vicinity of) the genes (http://ctglab.nl/software/MAGMA/ref_data/g1000_ceu.zip). Gene-wide p -values were converted to z -values reflecting the strength of the association of each gene with the phenotype (ICV), with higher z -values corresponding to stronger associations. Subsequently, we tested, whether the genes in the mTOR gene-set were jointly associated with ICV in the ENIGMA2 data set, using self-contained and competitive testing.²⁵⁷ For the gene-set analyses, we used an intercept-only linear regression model including a subvector corresponding to the genes in the gene-set. This self-contained analysis evaluating, whether the regression coefficient of this regression was larger than 0, tests whether the gene-set shows any association with ICV at all. Next, we tested whether genes in each gene-set were more strongly associated with ICV than all other genes in the

genome. Therefore, the regression model was then expanded including all genes outside the gene-set. With this competitive test, the differences between the association of the mTOR gene-set to genes outside this gene-set is tested, accounting for the polygenic nature of a complex trait like ICV. To account for the potentially confounding factors of gene size and gene density, both variables as well as their logarithms were included as covariates in the competitive gene-set analysis. Since self-contained tests do not take into account the overall level of association across the genome, gene-size (number of principal components, or SNPs) and gene density, we were interested in the competitive test for the current analysis. The same procedure was followed for analysis of the CHARGE cohort. In addition to the gene-set analyses within the individual cohorts, we meta-analyzed data of both cohorts on the gene-level followed by gene-set analysis. Post hoc, the potential effects of the two separate mTOR pathways in the gene-set (the PI3K-AKT-mTOR pathway (60 genes) and the RAS-MAPK-mTOR pathway (76 genes)) as well as the individual genes were investigated, by reviewing their gene test-statistics. Moreover, the Reactome_PI3K_AKT_activation gene-set, consisting of 38 genes, was tested for its association with ICV (downloaded from <http://software.broadinstitute.org/gsea/msigdb/genesets.jsp>). Genes were considered gene-wide significant, if they reached the Bonferroni correction threshold adjusted for the number of genes within the total gene-set ($N=96$; $p<0.000521$).

Generation of zebrafish *rheb* mutants

All animal experiments were carried out with the approval of the Institutional Animal Care and Use Committee (IACUC). Guide RNAs targeting the *Danio rerio* coding region of *rheb* were generated as described.¹⁸⁵ Subsequently, *rheb* guide oligonucleotide sequences (*rheb_ex3_g1F*: 5'-TAGGGTCGTGGAACGCAGCGTTCA-3' and *rheb_ex3_g1R*: 5'-AAACTGAACGCTGCGTTCCACGAC-3') were ligated into the pT7Cas9sgRNA vector (Addgene) into *Bsm* BI sites. For the generation of gRNA, the template DNA was linearized with *Bam* HI, purified by phenol/chloroform extraction and in vitro transcribed using the MEGAshortscript T7 kit (Invitrogen). To generate F0 CRISPR mutants we injected 1nl containing 100pg *rheb* guide RNA and 200ng Cas9 protein (PNA bio, CP01) to 1-cell stage embryos. To determine the efficiency of the guide RNA, embryos were allowed to grow to 5 days post fertilization (dpf), at which time they were killed and subjected to digestion with proteinase K (Life Technologies) to extract genomic DNA. The targeted locus was PCR amplified using the *drrehb_g1test_1 F* 5'-GAGTGATCAGCTGTGAAGAAGG-3' and *drrehb_g1test_1 R* 5'-GAACAGCGACAGGAGCTACA-3' primer pair. PCR amplicons were digested using T7 endonuclease I (New England Biolabs) at 37°C for 1h and were visualized on a 2% agarose gel. For Sanger sequencing of individual products from the *rheb* locus, PCR fragments from four embryos with a positive T7 assay were cloned into the pCR4/TOPO TA cloning vector (Life Technologies), and 40 colonies from each cloned embryo were Sanger sequenced. We observed sequence aberrations in ~75% of the evaluated *rheb* clones.

In vivo modeling in zebrafish embryos

The human wild-type²⁵⁸ mRNA of *RHEB* (NM_005614) was cloned into the pCS2+ vector and transcribed in vitro using the SP6 Message Machine kit (Ambion). The variants identified in *RHEB* in our patient cohort (RHEBp.P37L, RHEBp.S68P) were introduced using Phusion high-fidelity DNA polymerase (New England Biolabs) and custom-designed primers. We injected 50pg of WT or mutant RNA into wild-type zebrafish embryos at the 1- to 4-cell stage. For the experiments with rapamycin treatment we added 2.7nM of ready-made rapamycin solution in DMSO (R8781, Sigma-Aldrich) in each of the injection cocktails. For the headsize assay, the injected larvae were grown to 5 dpf and imaged live on dorsal view. The area of the head was traced excluding the eyes from the measurements and statistical significance was calculated using Student's *t*-test. All experiments were repeated three times and scored blind to injection cocktail.

Generation of constructs for mouse studies

The cDNA sequences from human *RHEB*-WT (NM_005614), and the variants found in the patient cohort (RHEBp.P37L and RHEBp.S68P) were synthesized by GeneCust, and cloned into our dual promoter expression vector. The dual promoter expression vector was generated from the pCMV-tdTomato vector (Clontech), in which the CMV promoter was replaced with a CAGG promoter followed by a multiple cloning site (MCS) and transcription terminator sequence. To assure expression of the *tdTOMATO* independent from the gene of interest, a PGK promoter was inserted in front of the tdTomato sequence (for a schematic overview of the expression vector see Supplemental Figure 7.3). For all the in vivo and in vitro experiments, the vector without a gene inserted in the MCS was used as control (control vector).

Mice used for the in vitro and in vivo studies

For the neuronal cultures, FvB/NHsD females were crossed with FvB/NHsD males (both ordered at 8–10 weeks old from Envigo). For the in utero electroporation female FvB/NHsD (Envigo) were crossed with male C57Bl6/J (ordered at 8–10 weeks old from Charles River). All mice were kept group-housed in IVC cages (Sealsafe 1145T, Tecniplast) with bedding material (Lignocel BK 8/15 from Rettenmayer) on a 12/12h light/dark cycle in 21°C (± 1 °C), humidity at 40–70% and with food pellets (801727CRM(P) from Special Dietary Service) and water available ad libitum. All animal experiments were approved by the Erasmus MC institutional Animal Care and Ethical Committee, in accordance with European and Institutional Animal Care and Use Committee guidelines.

In vitro modeling in mouse primary hippocampal neurons

Primary hippocampal neuronal cultures were prepared from FvB/NHsD wild-type mice according to the procedure described in Goslin and Banker.²⁵⁹ Briefly, hippocampi were isolated from brains of E16.5 embryos and collected altogether in 10ml of neurobasal

medium (NB, Gibco) on ice. After two washes with NB, the samples were incubated in pre-warmed trypsin/EDTA solution (Invitrogen) at 37°C for 20min. After two washes in pre-warmed NB, the cells were resuspended in 1.5ml NB medium supplemented with 2% B27, 1% penicillin/streptomycin and 1% glutamax (Invitrogen), and dissociated using a 5ml pipette. Following dissociation, neurons were plated on poly-D-lysine (25mg/ml, Sigma) coated 15 mm glass coverslips at a density of 3×10^4 or 5×10^4 cells per coverslip for the axon length measurements and 1×10^6 cells per coverslip for all the other experiments. The plates were stored at 37°C/5% CO₂ until the day of transfection. Neurons were transfected at 3 days in vitro (DIV3, DIV7, and DIV14) with the following DNA constructs: control vector (1.8µg per coverslip), RHEB-WT, RHEBp.P37L, and RHEBp.S68P (all 2.5µg per coverslip). Plasmids were transfected using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). For the rescue experiments, 20nM Rapamycin (dissolved in 0.01% Ethanol) was applied to the culture 1 day post transfection for 3 days. Neurons were fixed 4 days (Rapamycin experiment) or 5 days (for soma size) post transfection with 4% paraformaldehyde (PFA)/4% sucrose and incubated overnight at 4°C with primary antibodies in GDB buffer (0.2% BSA, 0.8M NaCl, 0.5% Triton X-100, 30mM phosphate buffer, pH 7.4). The following primary antibodies were used: guinea-pig anti MAP2 (1:500, catalogue number: 188004, Synaptic System) to stain for dendrites, and rabbit anti-RHEB (1:100, catalogue number: 4935, Cell Signaling). Donkey anti-guinea-pig-Alexa647- and donkey anti-rabbit-Alexa488-conjugated were used as secondary antibodies (all 1:200, catalogue numbers: 706-605-148 and 711-545-152, respectively, Jackson ImmunoResearch). Slides were mounted using mowiol-DABCO mounting medium. Confocal images were acquired using a LSM700 confocal microscope (Zeiss). For the analysis of the neuronal transfections, at least ten distinct confocal images ($\times 20$ objective, 0.5 zoom, 1024×1024 pixels; neurons were identified by the red immunostaining signal) were taken from each coverslip for each experiment. ImageJ software was used for the analysis of the soma size, by drawing a line around the soma of the cell. For each coverslip, the area of the transfected cells was normalized against the area of the non-transfected cells (five cells per coverslips). These values were then normalized against the mean value of the control (control vector).

In vivo modeling in mice

The in utero electroporation was performed as described before.²⁶⁰ Pregnant FvB/NHsD mice at E14.5 of gestation were used to target the progenitor cells giving rise to pyramidal cells of the layer 2/3.²⁶¹ The DNA construct (1.5–3µg/µl) was diluted in fast green (0.05%) and injected in the lateral ventricle of the embryos while still in utero, using a glass pipette controlled by a Picospritzer® III device. To ensure proper electroporation of the injected DNA constructs (1–2µl) into the progenitor cells, five electrical square pulses of 45V with a duration of 50ms per pulse and 150ms inter-pulse interval were delivered using tweezer-type electrodes connected to a pulse generator (ECM 830, BTX Harvard Apparatus). The positive pole was targeting the developing somatosensory cortex. The following plasmids were

injected: control vector, *RHEB*-WT, *RHEB*p.P37L and *RHEB*p.S68P. After birth, pups were sacrificed at P0 or P7 for histochemical processing (to investigate neuronal migration) or used to monitor seizure development.

For the migration analysis, confocal images ($\times 10$ objective, 0.5 zoom, 1024×1024 pixels) were taken from 2 to 3 non-consecutive sections from 2 and 3 electroporated animals per control and *RHEB*-containing plasmids, respectively. Images were rotated to correctly position the cortical layers, and the number of cells in different layers were counted using ImageJ using the analyze particles plugin. The results were exported to a spreadsheet for further analysis. Cortical areas from the pia to the ventricle were divided in 10 equal-sized bins and the percentage of tdTOMATO-positive cells per bin was calculated.

For immunofluorescence, mice were deeply anesthetized with an overdose of Nembutal and transcardially perfused with 4% PFA. Brains were extracted and post-fixed in 4% PFA. Brains were then embedded in gelatin and cryoprotected in 30% sucrose in 0.1M phosphate buffer (PB), frozen on dry ice, and sectioned using a freezing microtome ($40/50\mu\text{m}$ thick). Free-floating coronal sections were washed in 0.1M PB and a few selected sections were counterstained with 4',6-diamidino-2-phenylindole solution (DAPI, 1:10,000, Invitrogen) before being mounted with mowiol® (Sigma-Aldrich) on glass. Overview images of the coronal sections were acquired by tile scan imaging using a LSM700 confocal microscope (Zeiss) with a $\times 10$ objective. Zoom-in images of the targeted area were taken using a $\times 20$ objective. For seizure observations, mice obtained after in utero electroporation were observed daily starting at P18. General behavior was observed by looking for abnormal behaviors such as hyperactivity, the presence of stereotypical behaviors and the presence of tonic-clonic seizures, either spontaneous or induced upon mild handling. Weaned mice were video-monitored for 24h per day in the *Phenotyper* (Noldus), to assess seizure onset. Abnormal behaviors and onset of seizures were scored and analysed for each mouse by an expert experimentalist who had been blinded to the identity of samples (i.e., which plasmid had been transfected).

Statistical analysis used for the mouse studies

Statistical difference in soma size between the *RHEB* WT and mutants was determined using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test for multiple comparisons. The effect of Rapamycin treatment on soma size was determined using two-way analysis of variance (ANOVA) followed by Bonferroni's post hoc test for multiple comparisons. For the analysis of the in utero electroporation data, a two-way ANOVA repeated measure was performed, followed by the Bonferroni's multiple comparisons test. For the analysis of epilepsy onset, the log-rank Mantel-Cox test was used. The significance threshold was set at $p < 0.05$. Data are presented as mean \pm standard error of the mean (SEM).



8

Recurrent *de novo* heterozygous mutations disturbing the GTP/GDP binding pocket of *RAB11B* cause intellectual disability and a distinctive brain phenotype

This chapter has been published as:

Ideke J.C. Lamers*, **Margot R.F. Reijnders***, Hanka Venselaar, Alison Kraus, DDD Study, Sandra Jansen, Bert B.A. de Vries, Gunnar Houge, Gyri Aasland Gradek, Jieun Seo, Murim Choi, Jong-Hee Chae, Ineke van der Burgt, Rolph Pfundt, Stef J.F. Letteboer, Sylvia E.C. van Beersum, Simone Dusseljee, Han G. Brunner, Dan Doherty, Tjitske Kleefstra*, Ronald Roepman*

American Journal of Human Genetics (2017) 101, 824-832

* These authors contributed equally

Abstract

The Rab GTPase family comprises ~70 GTP-binding proteins, functioning in vesicle formation, transport and fusion. They are activated by a conformational change induced by GTP-binding, allowing interactions with downstream effectors. Here, we report five individuals with two recurrent *de novo* missense mutations in *RAB11B*; c.64G>A; p.Val22Met in three individuals and c.202G>A; p.Ala68Thr in two individuals. An overlapping neurodevelopmental phenotype, including severe intellectual disability with absent speech, epilepsy, and hypotonia was observed in all affected individuals. Additionally, visual problems, musculoskeletal abnormalities, and microcephaly were present in the majority of cases. Re-evaluation of brain MRI images of four individuals showed a shared distinct brain phenotype, consisting of abnormal white matter (severely decreased volume and abnormal signal), thin corpus callosum, cerebellar vermis hypoplasia, optic nerve hypoplasia and mild ventriculomegaly. To compare the effects of both variants with known inactive GDP- and active GTP-bound RAB11B mutants, we modeled the variants on the three-dimensional protein structure and performed subcellular localization studies. We predicted that both variants alter the GTP/GDP binding pocket and show that they both have localization patterns similar to inactive RAB11B. Evaluation of their influence on the affinity of RAB11B to a series of binary interactors, both effectors and guanine nucleotide exchange factors (GEFs), showed induction of RAB11B binding to the GEF SH3BP5, again similar to inactive RAB11B. In conclusion, we report two recurrent dominant mutations in *RAB11B* leading to a neurodevelopmental syndrome, likely caused by altered GDP/GTP binding that inactivate the protein and induce GEF binding and protein mislocalization.

Report

The application of whole-exome sequencing (WES) as a genetic test for intellectual disability (ID) and developmental delay has increased the number of genetic causes to more than 1,500.^{32; 126} This has significantly enhanced our knowledge of molecular processes that regulate learning and memory as well as brain development, although the roles of many genes involved in these processes have not yet been defined. Using trio-based WES in diagnostic and research settings,^{32; 126} we identified two different heterozygous *de novo* missense mutations in *RAB11B* (GenBank: NM_004218.3; MIM: 604198), c.64G>A; p.Val22Met and c.202G>A; p.Ala68Thr, in five unrelated individuals with severe ID (Figures 8.1A and B; Table 8.1). Written consent was obtained from the legal guardians for all individuals and the study was given IRB approval. Following identification of the *RAB11B* mutations of the first two individuals at Radboud University Medical Center, Nijmegen (individual 1: p.Val22Met; individual 4: p.Ala68Thr), a query via GeneMatcher⁸⁸ so resulted in the identification of individual 2 at Haukeland University Hospital, Bergen, and individual 3 at Seoul National University College of Medicine, Seoul. Both individuals harbored the variant p.Val22Met, similar to individual 1. A fifth individual (individual 5; DECIPHER ID: 263643), reported before as part of a study sequencing a large ID cohort,³³ was identified through international collaboration. Interestingly, this individual had the same variant, p.Ala68Thr, as individual 4 (Figure 8.1B). Neither identified missense mutation was present in the ExAC database or in-house control databases and both variants affect highly conserved amino acids in 90% of Rab family members¹⁷⁶ and among different species (Figure 8.1B). Additionally, the ExAC database has shown that *RAB11B* is significantly depleted of rare missense variants (Z score 3.48) in healthy controls.⁷¹

RAB11B is a member of the large Rab GTPase protein subfamily of RAS GTPases, consisting of almost 70 small ~21 kDa monomeric GTP-binding proteins. They serve as molecular switches that function in vesicle formation, transport, tethering, and fusion.^{262; 263} The tightly regulated spatiotemporal activity of Rabs is controlled by guanine nucleotide exchange factors (GEFs) that catalyze the GDP/GTP-exchange and GTPase activating proteins, which catalyze the hydrolysis of GTP into GDP.²⁶⁴⁻²⁶⁶ Rabs mainly interact with downstream GTPase effector proteins in their GTP-bound active conformation.^{265; 267} Rab GTPases and their effector proteins have been described to play a role in neuronal development and the shaping of cognitive functions.²⁶⁸ Several genes within this Rab GTPase family have been associated with neurodevelopmental disorders and micro- or macrocephaly, for example *RAB18* [MIM: 602207] and *RAB3GAP1* [MIM: 602536] are associated with Warburg micro syndrome [MIM: 600118], and *RAB39B* [MIM: 300774] is associated with Waisman syndrome [MIM: 311510] and Mental Retardation, X-linked 72 [MIM: 300271].

All five individuals that carried *de novo* mutations in *RAB11B* showed severe ID with motor delay and absent speech. A variety of neurological problems were present in all

Table 8.1 Clinical features of individuals with *de novo* mutations in *RAB11B*

	Individual 1 (Nijmegen)	Individual 2 (Bergen)	Individual 3 (Seoul)	Individual 4 (Nijmegen)	Individual 5 (DDD)
Gender	Female	Female	Male	Male	Female
Age of examination	13 years	4.5 years	8 years, 5 months	11 years	8 years, 8 months
Mutation (NM_004218.3)					
Chromosome position (Hg19)	g.8464770G>A	g.8464770G>A	g.8464770G>A	g.8464908G>A	g.8464908G>A
cDNA change	c.64G>A	c.64G>A	c.64G>A	c.202G>A	c.202G>A
Amino acid change	p.(Val22Met)	p.(Val22Met)	p.(Val22Met)	p.(Ala68Thr)	p.(Ala68Thr)
Growth					
Height	152 cm (-1 SD)	-2 SD	112 cm (-1.9 SD)	141 cm (-1.5 SD)	121 cm (-1.5 SD)
Weight	47.8 kg (+1.5 SD)	NR	34.1 kg (+1.7 SD)	34.5 kg (+0.7 SD)	25.7 kg (-0.5 SD)
Head circumference	49 cm (-3 SD)	-4 SD	52 cm (+0.6 SD)	50 cm (-2.2 SD)	48 cm (-3 SD)
Development					
Intellectual disability	Yes – severe/profound	Yes	Yes	Yes – severe/profound	Yes- severe/profound
First words	Absent speech	Absent speech	Absent speech	Absent speech	Absent speech
First steps with support	3 years	3 years	6 years	Unknown	8 years
Neurological					
Epilepsy	Possibly – to investigate	No	Yes	Yes	Single generalized seizure
Hypotonia	Yes – childhood	Yes	Yes	Yes	No
Spasticity	NR	No	No	Yes	Yes
Dystonia	NR	No	No	Yes	Yes
Abnormal gait	Yes – ataxic; broad-based	NR	NR	Yes – ataxic; broad-based	Yes – broad-based
Nystagmus	Yes	NR	NR	No	Yes - horizontal
Ophthalmological abnormalities					
Refraction abnormality	Hypermetropia – Mild	NR	No	Hypermetropia	No
Strabismus	Yes	NR	No	No	Yes
Other	Delayed visual maturation	Reduced vision	None	None	Optic atrophy
Musculoskeletal abnormalities					
Developmental hip dysplasia	Yes – mild, non-progressive	Yes – right-sided, non-progressive	Yes	No	Yes – requiring surgery
Tapering fingers	Yes	Yes	Yes	Yes	Yes
Other	Pes cavus; shortened achilles tendons; prominent steloideus ulnae	2 cm anisomelia; adducted thumbs; bilateral club foot	None	No	Long fingers

Other					
	Drooling; Simean crease; neonatal feeding difficulties	Bilateral palsy nervus laryngeus recurrens, Diabetes Mellitus Type 1, hydrocephalus	Acanthotic skin, Epidermal nevus in face, neck, trunk; Short neck; Obstructive sleep apnea; cryptorchidism	Simean crease	Bruxism.

Abbreviations: NR, Not Reported; SD, Standard Deviation.

affected individuals, including hypotonia (4/5), epilepsy (3/5), spasticity (2/4), dystonia (2/4), broad-based (3/3), and ataxic gait (2/3) and nystagmus (2/3) (Table 8.1). Re-evaluation of brain MRI images of four individuals showed similar brain imaging abnormalities (Figure 8.2 and Supplemental Table 8.1). All had severely decreased white matter volume (cerebral cortex more severely affected than cerebellum), thin corpus callosum, and mild ex vacuo lateral ventriculomegaly affecting the frontal horns and body of the ventricles more than the occipital and temporal horns. Other features that were scorable in a subset of images included cerebellar vermis hypoplasia (3/3), thin brainstem (3/3), early global white matter signal abnormalities consistent with delayed myelination (2/2), later patchy white matter signal abnormalities consistent with injury (4/4), optic nerve hypoplasia (2/2), and what appears to be atypical partial rhombencephalosynapsis (1/3). Although two other affected individuals did not have rhombencephalosynapsis (partial or complete absence of the cerebellar vermis with fusion of the cerebellar hemispheres), the width of the cerebellar vermis was subjectively narrower than typical. Besides the neurological phenotype, ophthalmological and musculoskeletal abnormalities were present in the majority of individuals. Microcephaly was observed in three individuals (Table 8.1).

To study the effects of the identified mutations on the well described three-dimensional protein structure of RAB11B, we modeled the substitutions on Protein Data Bank entry PDB: 2F9M²⁶⁹ using YASARA software,²⁷⁰ and compared to the described active, GTP-locked RAB11B mutant p.Gln70Leu and inactive, GDP-locked RAB11B mutant p.Ser25Asn (Figure 8.1C).²⁷¹ The p.Ser25Asn variant was shown to disrupt the binding of a magnesium molecule which is essential for GTP binding, consequently locking the GTPase in an inactive GDP bound, non-membrane-associated state.²⁷²⁻²⁷⁴ In contrast, the p.Gln70Leu variant affects a conserved residue in the flexible switch II region, which is essential for catalysis, and therefore this variant is predicted to fix the GTPase in a GTP-bound state, which is constitutively active (Figures 8.1B and C).^{274; 275} Both mutations identified in this study are situated in close proximity to the binding pocket for GTP/GDP and specifically on the side where the phosphate groups of GTP/GDP are positioned (Figure 8.1C). The valine residue at position 22 is located at the base in the middle of the GTP/GDP binding pocket, thereby being responsible for its shape (Figures 8.1B and D). Substitution

into methionine brings in a larger amino acid which causes a predicted reorganization of the region around this residue, altering the shape of the binding pocket. This likely disrupts binding of GDP and GTP, resulting in a nucleotide-free, inactive state of RAB11B. The second mutation affects an alanine at position 68 and is located within a flexible loop of the switch II region (Figures 8.1B and E), which together with the switch I region specifically interacts with effector proteins when GTP is bound.²⁶⁷ This flexible loop provides more space for a bigger amino acid residue, however, the mutant residue is positioned closely to the magnesium molecule, as well as the phosphate groups of GTP/GDP, and thereby alters the binding pocket and is predicted to disrupt nucleotide binding. Furthermore, the location within the switch region suggests that the mutation causes altered binding with effector proteins. Taken together, both identified mutations are predicted to alter the GTP/GDP binding pocket of RAB11B, hence it is likely that the functionality of the protein is affected.

RAB11B is one of three genes in the Rab11 GTPase subfamily, together with *RAB11A* [MIM: 605570] and *RAB25* [MIM: 612942], and this subfamily specifically associates with recycling endosomes.^{262; 276} *RAB11A* is the best-characterized family member and is ubiquitously expressed,²⁷⁷ in contrast to specific expression of *RAB25* and *RAB11B* in epithelial tissue²⁷⁸ or in heart, testis, and brain,²⁷⁹ respectively. The proteins have been implicated in the regulation of vesicular trafficking between the recycling endosome compartment and early endosomes to the trans Golgi network and plasma membrane.^{272;}

²⁸⁰ Furthermore, it has been shown that RAB11 is essential for ciliogenesis,²⁸¹⁻²⁸³ and that the proteins are localized to peri-centrosomal recycling endosomes concentrated at the base of the cilium.²⁸⁴

To evaluate whether the identified mutations affect the sub-cellular localization of RAB11B, potentially at the cilium, we transfected human TERT-immortalized retinal pigment epithelium 1 (hTERT RPE1) cells with cDNA expression constructs encoding WT (UniProt: Q15907) or mutant RAB11B, fused to an N-terminal 3xFLAG tag. A marker for the peri-centrosomal region (Pericentriolar Material 1; PCM1) was used to assess the effect on the peri-centrosomal localization (Figures 8.3A–F), and the marker acetylated tubulin to assess cilium morphology (Supplemental Figure 8.1). Wild-type RAB11B localization was scattered throughout the cell with a punctuated pattern as shown in Figure 8.3A. For some cells, puncta were enriched, as expected, at the peri-centrosomal region. We observed that the constitutively active mutant RAB11B-p.Gln70Leu had a strong organization in puncta throughout the cytoplasm and a consistent localization at the peri-centrosomal region in all cells (Figure 8.3B). In contrast, GDP-bound inactive RAB11B-p.Ser25Asn showed a generally dispersed cytosolic localization, losing its association with the peri-centrosomal region, but with enrichment near the nucleus suggesting Golgi localization (Figure 8.3C). This was confirmed using Golgi marker GM130 (Figure 8.3G–L). Our results upon expression of recombinant WT, p.Ser25Asn, and p.Gln70Leu RAB11B confirmed earlier studies

on the effect of these variants on the localization of this GTPase.^{272; 273} Localization of both RAB11B-p.Val22Met (Figure 8.3D) and RAB11B-p.Ala68Thr (Figure 8.3E) showed homogeneous localization in the cytoplasm, no association with the peri-centrosomal

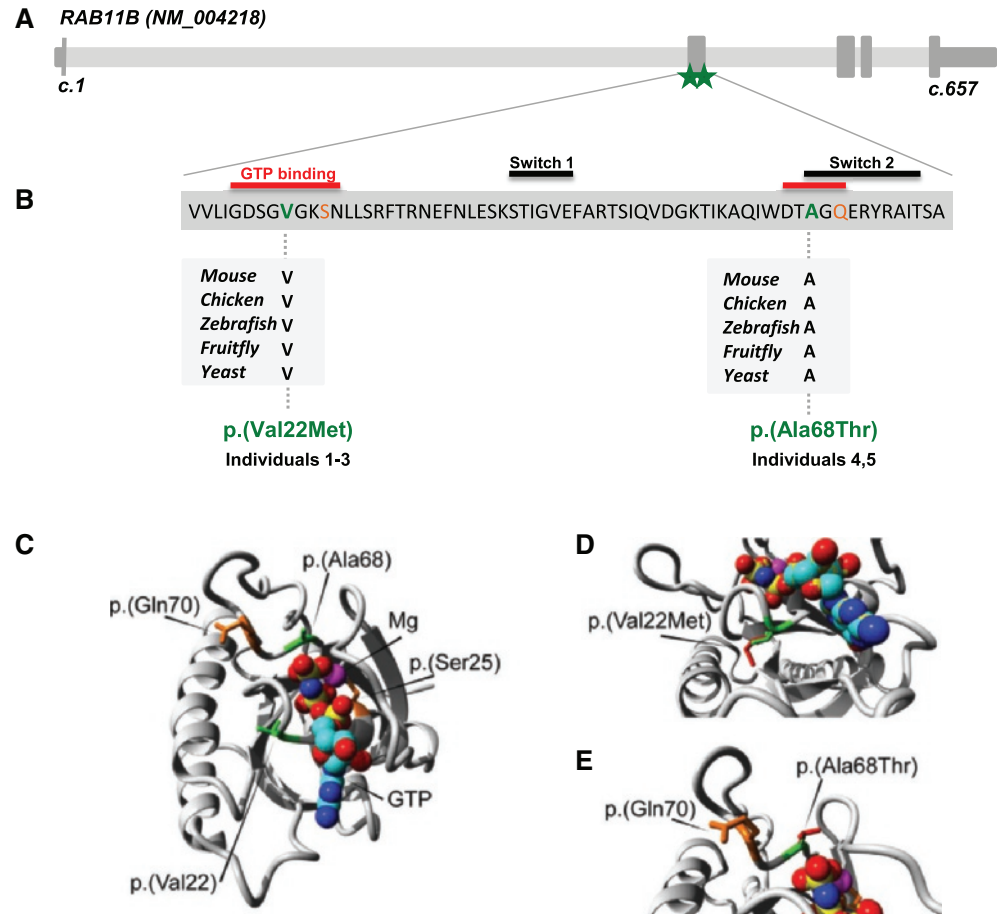


Figure 8.1: RAB11B Mutations in five individuals with severe ID.

(A) cDNA composition of RAB11B and location of exons (dark gray) and identified mutations (green stars). (B) Amino acid sequence of exon 2 of RAB11B. Both identified variants (green) localize at one of the two GTP binding sites (red stripes) and p.Ala68Thr also involves the switch 2 region (black stripe). Both identified mutations affect highly conserved amino acid residues among different species (light gray box). The known inactive variant p.Ser25Asn and known active variant p.Gln70Leu are marked orange. (C–E) Structural characterization of mutations. In all panels, the overall Protein DataBank structure 2f9m of RAB11B²⁶⁹ is colored gray and shown in ribbon representation, GNP (phosphoaminophosphonic acid-guanylate ester, a non-hydrolyzable analog of GTP) is multicolored in a spacefill representation with the guanine ring in blue shades and the phosphates in red and yellow. The magnesium molecule is represented as a purple sphere. The side chains of the affected amino acids in the affected individuals are green and for the GDP/GTP-locked mutants are orange. (C) Overview of RAB11B bound to GNP and magnesium and the affected amino acids are highlighted. (D) Close up of the p.Val22Met variant shows the side chains of valine at position 22 in green and the methionine in red. (E) Close up of the p.Ala68Thr variant shows the side chains of alanine at position 68 in green and threonine in red, and the mutated glutamine at position 70 which is mutated in the GTP-locked mutant is in orange.

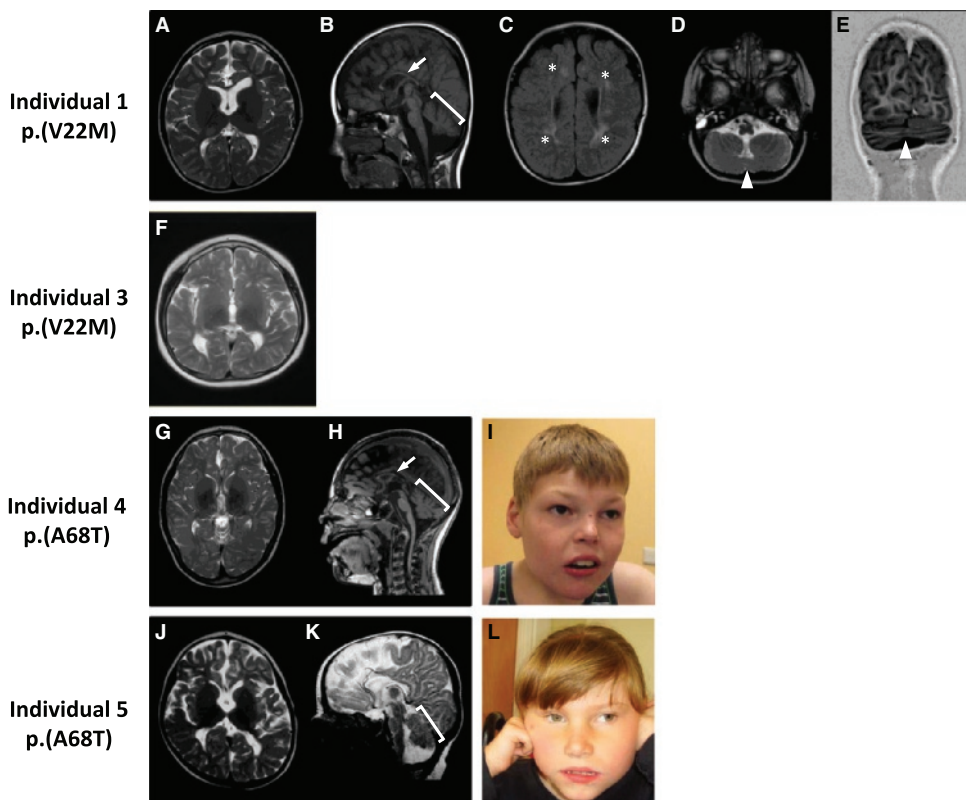


Figure 8.2: Brain imaging features and facial photographs of individuals with RAB11B-related intellectual disability.

(A, F, G, J) Markedly decreased cortical white matter volume in individuals 1, 3, 4, and 5. (B, H, K) Markedly thin corpus callosum (arrows), mildly thin brainstem, and mildly small, atrophic-appearing cerebellar vermis (bracket) in individual 1, 4, and 5. (C) Increased T2/FLAIR signal in the cortical white matter (asterisks) in individual 1. (D and E) Atypical partial rhombencephalosynapsis in individual 1 (arrowheads). (A, D, F, G, and H) are T2-weighted axial images. (B and H) are T1-weighted sagittal images. (C) is a T2/FLAIR-weighted axial image. (E) is a reformatted inversion-recovery coronal image. (K) is a T2-weighted sagittal image. (I and L) Facial photographs show upward slanted palpebral fissures, periorbital fullness, full nasal tip, and hypotonic face in individual 4 and upward slanted palpebral fissures, deep set eyes, and short philtrum in individual 5.

region and co-localization with Golgi marker GM130 (Figures 8.3J and K), similar to the GDP-bound inactive mutant p.Ser25Asn. Quantification of the co-localization patterns with PCM1 and GM130 confirmed our observations (Figure 8.3F and L). The morphology of the cilium appeared normal in cells transfected with each of the four mutant constructs (Supplemental Figure 8.1). We reproduced the localization patterns at the base of the cilium or at the Golgi for all RAB11B variants in a different cell line: mouse inner medullary collecting duct 3 (IMCD3) cells (Supplemental Figure 8.2). In short, introduction of either one of the identified RAB11B variants results in a similar localization pattern as the GDP-bound inactive form, without an obvious difference between the two. Interestingly, previous

studies showed that GDP-locked GTPases lose their membrane association.^{272; 274} For the p.Ser25Asn mutant of RAB11B, we indeed observe a localization pattern that is similar to that of the published RAB11B-ΔC mutant where five C-terminal amino acids were substituted to abolish prenylation, and consequently disturbed vesicular membrane association.²⁷³ The fact that both identified variants result in a comparable RAB11B localization pattern as the p.Ser25Asn mutant, suggests that in the affected individuals, the association of RAB11B with (vesicular) membranes is affected, contributing to the pathogenic effects of the identified variants.

To further assess the effect of the identified variants on the functionality of RAB11B, we screened for binary interaction partners of RAB11B by using a GAL4-based yeast two-hybrid screen of cDNA libraries from neuronal tissues (brain and retina) as previously described.²⁸⁵ In physiological conditions, RAB11B is either GDP or GTP bound. Therefore we used both p.Ser25Asn and p.Gln70Leu mutant constructs as baits to screen for potential interactors, which identified two (p.Ser25Asn) and eight (p.Gln70Leu) different potential interactors (Figure 8.4 and Supplemental Figure 8.3). All interactions were confirmed using independent co-transformation assays with validation of reporter gene activity (Supplemental Figure 8.3). The identified interactors were all previously reported as Ral/Rac or Rab11 specific interacting partners (Supplemental Table 8.2), validating our assay. Interestingly, three direct interactors identified in our yeast screen are encoded by genes where loss of function mutations are associated with neurodevelopmental disorders: *CNKSR2* [MIM: 300724] with X-linked Intellectual Disability,²⁸⁶⁻²⁸⁸ *MYO5A* [MIM: 160777] with Griscelli syndrome type 1 [MIM: 214450], and *TRAPPC9* [MIM: 611966] with Recessive Mental Retardation 13 [MIM: 613192] (Table S2).

To assess whether the identified variants affect the affinity of RAB11B to any of the interactors, we co-expressed p.Val22Met and p.Ala68Thr mutant constructs with each interactor and compared the binding affinities semiquantitatively by evaluation of the reporter gene-activation levels (Figure 8.4 and Supplemental Figure 8.3, and Supplemental Table 8.2). Wild-type, GDP-, and GTP-bound RAB11B were taken along as controls. Because GTP is more abundantly present in the cytosol than GDP, the WT RAB11B construct is expected to be predominantly GTP-bound. We indeed observed that the interaction pattern of the WT was comparable to the constitutively active mutant (Figure 8.4 and Supplemental Figure 8.3). When either one of the identified variants was introduced, the binding pattern was mostly similar to that of WT RAB11B, but one important difference could be distinguished: both mutant proteins were able to bind to the one bona fide GEF identified in our screen, SH3BP5, while WT RAB11 could not (Figure 8.4 and Supplemental Figure 8.3).

The affinity of the p.Val22Met mutant protein to SH3BP5 was strong and comparable to the GDP-locked mutant protein p.Ser25Asn, while the affinity of the p.Ala68Thr mutant protein

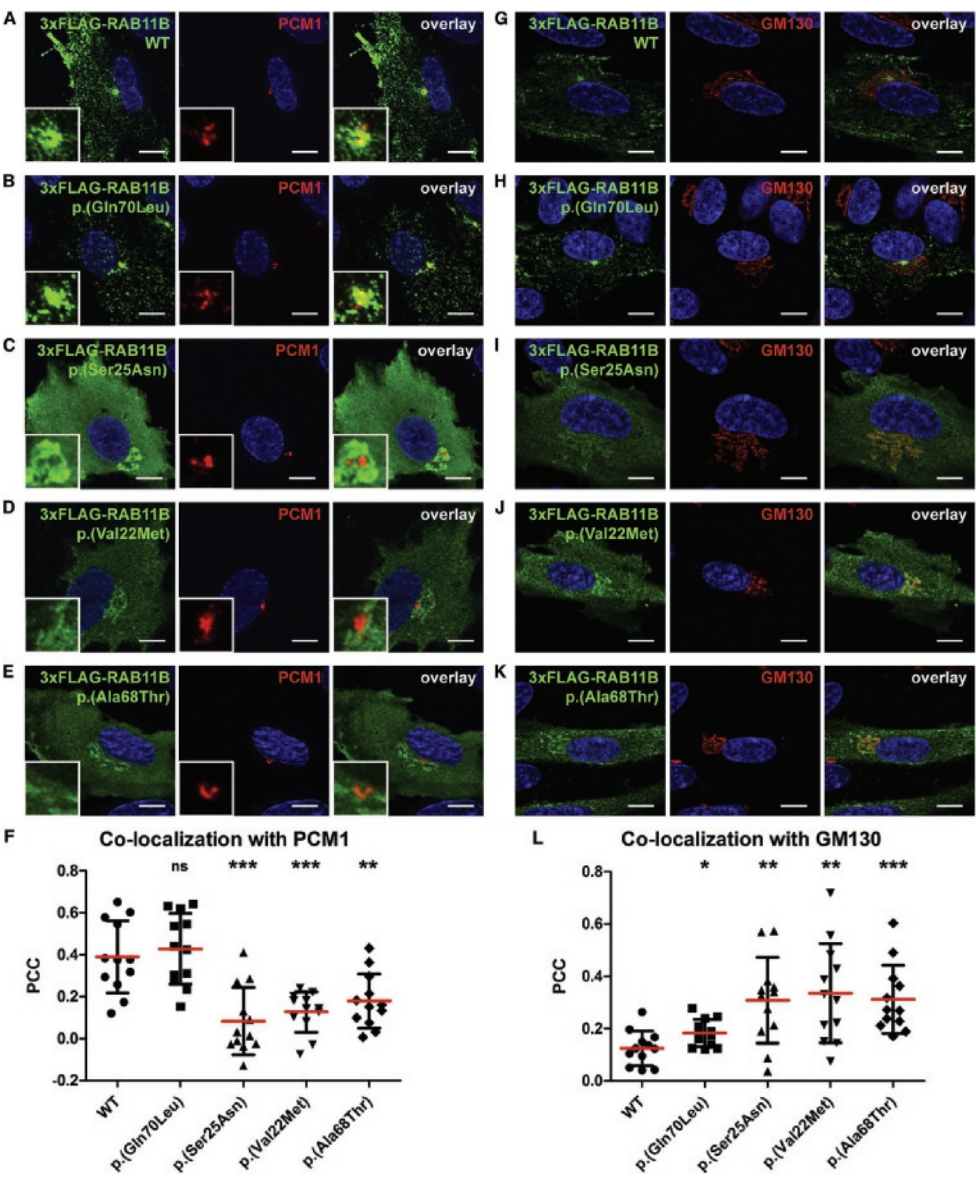


Figure 8.3: Localization of wild-type and mutant RAB11B in hTERT-RPE1 cells. Shown in green is the expression of recombinant 3xFLAG-RAB11B wild-type (A and G) and 3xFLAG-RAB11B variants p.Gln70Leu (B and H), p.Ser25Asn (C and I), p.Val22Met (D and J), p.Ala68Thr (E and K) in hTERT RPE1 cells, detected using anti-FLAG antibodies (rabbit polyclonal, #F7425, Sigma Aldrich; dilution 1:500 in PBS). N-terminal tagging was performed to not disturb the prenylation of the C terminus of RAB11B, which is required for membrane association.^{265; 273} Recombinant protein expression, immunocytochemistry, and image capture was performed as described previously.²⁹⁹ All secondary antibodies were Alexa Fluor conjugates (Thermo Fisher Scientific). The following transfection efficiencies were obtained per 3xFLAG-RAB11B construct: wild-type, 14%; p.Gln70Leu, 19%; p.Ser25Asn, 9%; p.Val22Met, 10%; p.Ala68Thr, 14%. Staining of Pericentriolar Material 1 (PCM1) using anti-PCM1 antibodies (goat polyclonal, #SC-50164, Santa Cruz Biotechnology; dilution 1:250 in PBS) was used to mark the peri-centrosomal region (red) and magnifications are shown in the insets (A–E). Staining of GM130 using anti-GM130 antibodies (mouse monoclonal, #610822, BD Biosciences; dilution 1:250 in PBS) was used to mark the Golgi

apparatus (in red; G-K). Co-localization with PCM1 (F) or GM130 (L) was quantified by calculating the Pearson's correlation coefficient (PCC) using the JACoP plugin in ImageJ ($N = 12$ cells for each construct, mean in red and error bars represent the SD).³⁰⁰⁻³⁰² The significance of difference with WT was calculated with a Student's *t* test (ns: not significant ($p > 0.05$); *: $p < 0.05$; **: $p < 0.005$; ***: $p < 0.0005$). Wild-type and GTP-bound active RAB11B (p.Gln70Leu) are showing punctuated localization with an enrichment at the peri-centrosomal region. Both identified variants p.Val22Met and p.Ala68Thr show similar localization as GDP-bound inactive RAB11B (p.Ser25Asn) with dispersed localization throughout the cytoplasm and enrichment at the Golgi apparatus. In all pictures, nuclei were stained with DAPI (blue). Scale bars represent 10 μ m.

was somewhat lower. Based on the 3D modeling and our localization assay, we already suggested that the p.Val22Met variant results in a nucleotide-free state by the inability of the protein to bind GDP or GTP. The interaction data support our hypothesis that variant p.Val22Met causes a nucleotide-free state, because GEFs are required for the stabilization of a GTPase if not bound to GTP or GDP.^{266; 274} A somatic substitution affecting the adjacent conserved glycine residue (RAB11B position 23; Supplemental Figure 8.4) has been described in RAS-GTPase RHOA (p.Gly17Val),²⁸⁹⁻²⁹¹ resulting in a nucleotide-free state of RHOA. As a result, mutant RHOA acts in a dominant-negative manner because it sequesters GEFs, which prohibits that these GEFs are available to activate wild-type RHOA.²⁹² With (1) the position of the p.Val22Met variant adjacent to this reported *RHOA* substitution, (2) the 3D modeling suggesting a nucleotide-free state of RAB11B (Figure 8.1C), and (3) the strong affinity of the mutant to bind the RAB11B GEF SH3BP5 (Figure 8.4 and Supplemental Figure 8.3), it is likely that the identified p.Val22Met variant acts in a dominant-negative manner as described for RHOA. Interestingly, we observed that this mutant RAB11B is still able to interact with effectors as well as GEFs (Figure 8.4 and Supplemental Figure 8.3). Also, for the p.Ala68Thr variant, the binding to SH3BP5 is much less pronounced and only detectable under less stringent assay conditions (-LWH + 5mM 3AT, Figure 8.4 and Supplemental Figure 8.3). It has been described that a specific dominant-negative mutant of Ras (Ras15A), member of the RAS-GTPase family as well, has stronger affinities for GEFs with more defective nucleotide binding, compared to another dominant-negative mutant of Ras (Ras17N) where a neighboring residue is affected.²⁹² This highlights the possibility of variable effects between variants in the GTP/GDP binding pocket on the affinity for GEFs and nucleotides. We argue that, despite these slight differences in affinity, both variants have similar consequences on RAB11B function, since individuals who harbor the p.Ala68Thr variant display strong phenotypic overlap with individuals carrying the p.Val22Met variant. Furthermore, both variants caused a similarly disturbed localization (Figure 8.3). Therefore, we hypothesize that p.Ala68Thr could act in a dominant-negative manner as well, with the same phenotypic consequences as the p.Val22Met variant.

The second GEF for RAB11B identified in our screen is TRAPPC9, member of the TRAPP II complex. This group of proteins is involved in intracellular membrane trafficking processes^{265; 293} and acts as a GEF for RAB11B orthologs in yeast.²⁹⁴⁻²⁹⁶ The fact that the p.Val22Met and p.Ala68Thr mutant proteins are not able to associate with TRAPPC9 seems to contradict

the results with the GEF SH3BP5. However, SH3BP5 was shown to be a bona fide GEF for RAB11,²⁹⁷ while TRAPPC9 requires the other subunits of the TRAPP II complex to act as a GEF. As the identified *RAB11B* mutations do not induce the TRAPPC9 binding, our data also suggest that TRAPPC9, as a single protein, does not act as a RAB11 GEF in contrast to SH3BP5.

It is likely that RAB11B disruption also has molecular consequences other than disturbed protein interactions that might alter several cellular mechanisms such as calcium influx, synaptic function, and neuronal migration,²⁶⁸ that could contribute to the underlying pathology. One explanation could be found in the mislocalization of RAB11B mutant proteins, which were able to bind effector proteins based on our yeast two-hybrid data. As a result of the membrane unbound situation suggested by the localization data, mislocalization of mutant RAB11B might therefore result in mislocalization of bound effectors, sequestering them from their usual sites of function.²⁹² Our data also show that RAB11B localization is altered at the basal body of the cilium (Supplemental Figure 8.1), and RAB11B is a direct interactor of the ciliary RAB Rabin8 (Figure 8.4 and Supplemental Table 8.2),²⁹⁸ although we observed no morphological changes of the cilium in hTERT-RPE1 cells. More subtle changes in cilium morphology could have been missed in our assay or cilia might have an altered signaling function without morphological abnormalities. Although speculative at this point without an animal model, cilia could be mainly affected in brain tissue, given the predominant expression of RAB11B in the brain.^{268; 279}

	Interactor	type of interactor	selection on -LWH + 5mM 3AT														
			p.(Val22Met)			p.(Ala68Thr)			p.(Ser25Asn)			p.(Gln70Leu)			WT		
			growth	α-gal	β-gal	growth	α-gal	β-gal	growth	α-gal	β-gal	growth	α-gal	β-gal	growth	α-gal	β-gal
fetal brain	SH3BP5	GEF	3	3	3	2	1	1	3	3	3	0	0	0	0	0	0
	TRAPPC9	GEF subunit	0	0	0	0	0	0	2	1	2	0	0	0	0	0	0
	CNKS2	Rac regulator	0	0	0	0	0	0	0	0	0	2	1	2	0	0	0
	RAB11FIP3	adapter/effector	3	3	3	2	0	2	0	0	0	3	3	3	2	0	2
bovine retina	Myosin-Va	effector	2	0	3	2	0	3	0	0	0	2	1	3	2	0	2
	RAB11FIP2	adapter/effector	3	3	3	3	3	3	1	0	3	3	3	3	3	3	3
	RAB11FIP3	adapter/effector	2	2	3	2	0	3	0	0	0	2	2	3	2	0	3
	RAB11FIP4	adapter/effector	2	0	2	1	0	0	0	0	0	2	0	2	0	0	0
human retina	Rabin8	effector	3	2	3	2	1	3	2	1	3	3	2	3	3	1	2
	Myosin-Va	effector	2	1	3	2	0	2	0	0	0	3	1	3	3	0	1
	Myosin-Vb	effector	2	2	3	2	1	3	0	0	0	3	3	3	3	1	3
	RAB11FIP3	adapter/effector	2	1	3	2	0	2	0	0	0	2	1	3	3	0	2
human retina	RAB11FIP5	adapter/effector	2	2	3	3	2	2	0	0	1	3	3	3	2	2	3
	Rabin8	effector	3	1	3	3	1	2	3	1	2	3	2	3	3	0	2

Figure 8.4: Identified interactors of RAB11B using yeast two-hybrid cDNA library screening and their affinity for RAB11B variants.

Co-transformations were performed in PJ69-4a yeast strains with the identified clones from the library screens fused to pAD together with WT or mutant RAB11B constructs fused to pBD. The clones are sorted according to the library in which they have been identified, as indicated in the far left column. The quantification of the selection of the growth of two-hybrid clones grown on medium lacking leucine, tryptophan and histidine with 3mM 3AT is quantified in the "growth" column. α-Galactosidase reporter gene activation (α-gal column) or β-galactosidase reporter gene activation (β-gal column) are quantified as well. Quantifications are on a scale from 0-3; with 0 for no reporter gene activation and 3 for highest reporter gene activation. Details of the identified clones in the cDNA library screens can be found in Table S2, including quantification of reporter gene activation on -LWHA medium. Original images are shown in Figure S2.

In conclusion, we identified two recurrent *de novo* missense mutations in *RAB11B* in five unrelated individuals with severe ID and specific brain abnormalities. We show that (1) both mutations affect the GDP/GTP binding pocket of this small GTPase, (2) the association of the mutant proteins with (vesicular) membranes is affected in localization studies, and (3) both mutations have limited effect on RAB11B effector protein binding, but can enhance the affinity to the GEF SH3BP5. We propose that these effects together cause distinct defects in several neuronal developmental processes, a combination that results in a neurodevelopmental syndrome in human.



9

***De novo* loss-of-function mutations in *USP9X* cause a female-specific recognizable syndrome with developmental delay and congenital malformations**

This chapter has been published as:

Margot R.F. Reijnders*, Vasilios Zachariadis*, Brooke Latour*, Lachlan Jolly*, Grazia M. Mancini, Rolph Pfundt, Ka Man Wu, Conny M.A. van Ravenswaaij-Arts, Hermine E. Veenstra-Knol, Britt-Marie M. Anderlid, Stephen A. Wood, Sau Wai Cheung, Angela Barnicoat, Frank Probst, Pilar Magoulas, Alice S. Brooks, Helena Malmgren, Arja Harila-Saari, Carlo M. Marcelis, Maaïke Vreeburg, Emma Hobson, V. Reid Sutton, Zornitza Stark, Julie Vogt, Nicola Cooper, Jiin Ying LIM, Sue Price, Angeline Hwei Meeng LAI, Deepti Domingo, Bruno Reversade, the DDD study, Jozef Gecz, Christian Gilissen, Han G. Brunner, Usha Kini*, Ronald Roepman*, Ann Nordgren*, Tjitske Kleefstra*

American Journal of Human Genetics (2016) 98, 373-381

* These authors contributed equally

Abstract

Mutations in more than a hundred genes have been reported to cause X-linked recessive intellectual disability (ID) mainly in males. In contrast, the number of identified X-linked genes in which de novo mutations specifically cause ID in females is limited. Here, we report 17 females with de novo loss-of-function mutations in *USP9X*, encoding a highly conserved deubiquitinating enzyme. The females in our study have a specific phenotype that includes ID/developmental delay (DD), characteristic facial features, short stature, and distinct congenital malformations comprising choanal atresia, anal abnormalities, post-axial polydactyly, heart defects, hypomastia, cleft palate/bifid uvula, progressive scoliosis, and structural brain abnormalities. Four females from our cohort were identified by targeted genetic testing because their phenotype was suggestive for *USP9X* mutations. In several females, pigment changes along Blaschko lines and body asymmetry were observed, which is probably related to differential (escape from) X-inactivation between tissues. Expression studies on both mRNA and protein level in affected-female-derived fibroblasts showed significant reduction of *USP9X* level, confirming the loss-of-function effect of the identified mutations. Given that some features of affected females are also reported in known ciliopathy syndromes, we examined the role of *USP9X* in the primary cilium and found that endogenous *USP9X* localizes along the length of the ciliary axoneme, indicating that its loss of function could indeed disrupt cilium-regulated processes. Absence of dysregulated ciliary parameters in affected female-derived fibroblasts, however, points toward spatiotemporal specificity of ciliary *USP9X* (dys-)function.

Report

X-linked intellectual disability (ID) with presumed recessive inheritance pattern is shown to be caused by mutations in more than a hundred genes^{303; 304} Most families display a clear X-linked segregation pattern, in which males are affected while females are unaffected or mildly affected carriers.³⁰⁵⁻³⁰⁷ In contrast, the number of identified X-linked genes in which *de novo* mutations cause ID specifically in females is limited.

Using whole-exome sequencing (WES), SNP array, array CGH, and CytoScan HD array in a diagnostic setting as described before,^{19; 28; 99; 308-310} we identified 13 *de novo* loss-of-function mutations in *USP9X* (Ubiquitin-specific protease 9 [MIM: 300072; GenBank: NM_001039590.2]) in females with ID/developmental delay (DD) and multiple congenital malformations (Figures 9.1A and B; Supplemental Table 9.1). Female 7 was previously reported as part of a large study sequencing individuals with ID, congenital anomalies, and/or autism with a targeted gene panel.³⁰⁹ Written consent was obtained from the legal guardians for all females and the study was given IRB approval. We recognized a similar pattern of facial characteristics, congenital malformations, and brain abnormalities in these females. Four additional affected females were identified because their phenotype was suggestive for *USP9X* mutations. Subsequently, *de novo* protein-truncating mutations and intragenic *USP9X* deletions were duly demonstrated by Sanger sequencing, WES, or CytoScan HD array (Figures 9.1A and B; Supplemental Table 9.1), illustrating the clinical recognizability of this new syndrome. All females (age ranging 2 years, 7 months to 23 years) with *de novo* mutations shared a distinct phenotype. They showed mild to moderate ID with motor and language delay, short stature, hearing loss, and distinct congenital malformations, notably choanal atresia, asymmetric hypomastia, cleft palate/bifid uvula, heart defects, progressive scoliosis, post-axial polydactyly, and anal abnormalities (Tables 9.1 and Supplemental Table 9.2; case studies in Supplemental Note). Shared facial characteristics included prominent forehead, low nasal bridge, prominent nose with flared alae nasi, thin upper lip, smooth and long philtrum, and ears that were low set, posteriorly rotated, and dysplastic (Figure 9.2A). In addition to the *USP9X* variant, female 5 also harbored a *de novo* variant in *PTPN11* (MIM: 176876), which has previously been reported to cause Noonan syndrome (MIM: 163950).³¹¹ Though all features that were observed in this female could potentially be explained by the *USP9X* variant itself (Supplemental Table 9.2), a contribution of aberrant *PTPN11* to phenotypic features such as intellectual disability, short stature, and heart defect in this female is likely. Neuroimaging reports were available for 13 out of 17 females (Supplemental Table 9.2). Detailed evaluation of brain images of five of these females (females 1, 2, 3, 7, and 16) showed asymmetric hypoplasia of the cerebellar vermis and hemisphere with a retrocerebellar cyst, short and thin corpus callosum, thin brainstem, and mildly abnormal frontal gyration pattern (Figure 9.3). Notably, we observed thyroid hormone abnormalities in six of the females, requiring medical treatment in three of them.

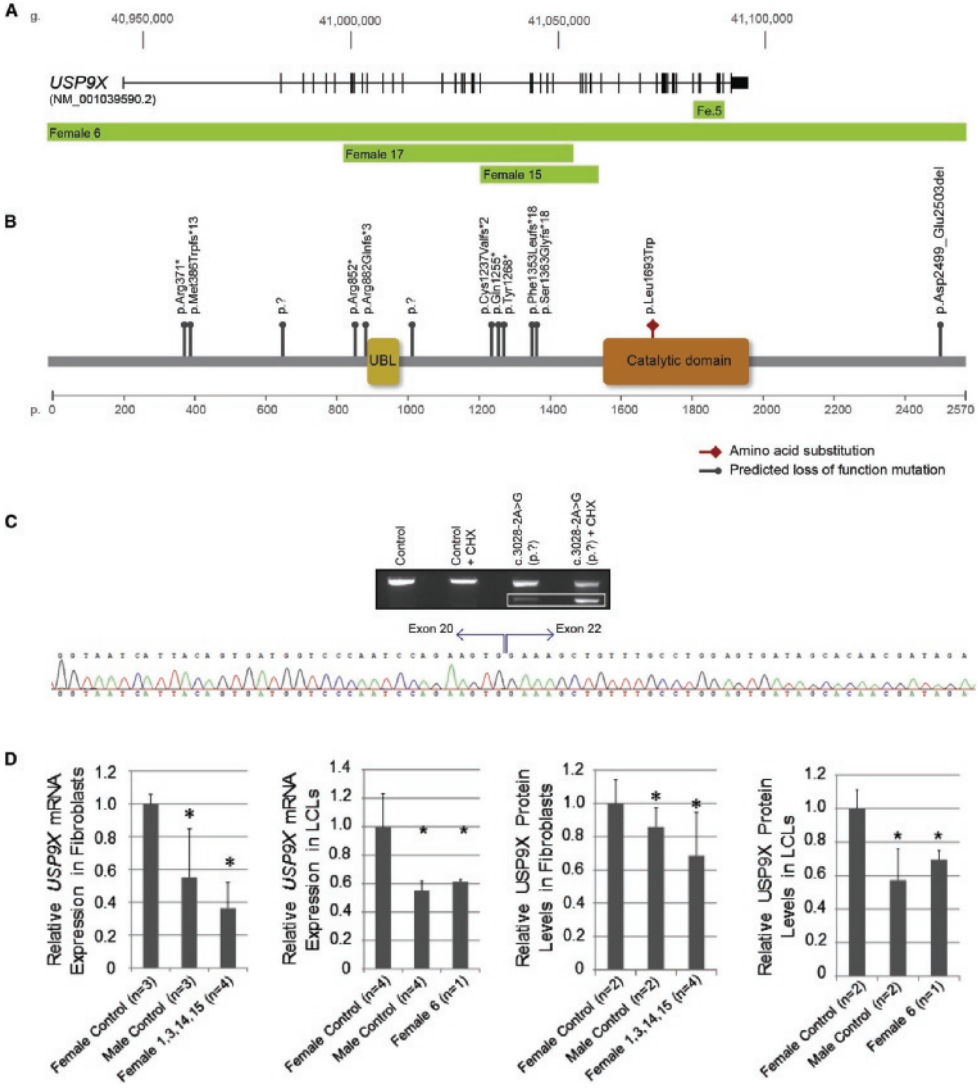


Figure 9.1: Identified de novo USP9X loss-of-function mutations.

(A) Detailed view of the USP9X (GenBank: NM_001039590.2) region and the reported deletions. (B) Overview of USP9X including UBL and catalytic domain and the location of reported mutations according to their relative position at the protein level. The reported amino acid substitution is located within the catalytic domain. (C) RNA was extracted from both control and affected female (c.3028-2A>G [p. (?)]) fibroblasts cultured under normal conditions or in the presence of cycloheximide (CHX) to inhibit NMD. After cDNA synthesis and PCR, agarose gel analysis showed two different product sizes generated from the c.3028-2A>G transcript but only one from the control fibroblast transcript. Excision and sequencing of the additional band revealed that the aberrant USP9X transcript lacked exon 21. The level of the aberrant transcript was increased 4-fold when fibroblasts were treated with cycloheximide, confirming that the aberrant transcript was indeed subjected to NMD. (D) USP9X expression is depleted in female cell lines harboring loss-of-function alleles. Relative qPCR analysis of USP9X mRNA and relative quantification of immunoblot analysis of USP9X protein derived from female and male control cell lines and from affected female cell lines. *n* = the number of individual cell lines analyzed. Each cell line analyzed in quadruplicate. Error bars represent SDs. Asterisk (*) indicates significantly different from female controls, $p < 0.05$ by Student's *t* test.

Table 9.1 Clinical features of females with *de novo* USP9X loss of function mutations

	Percentage	Number
Development		
Intellectual disability or developmental delay	100%	17/17
Growth		
Short stature	53%	9/17
Congenital abnormalities		
Eye abnormality	59%	10/17
Choanal atresia	35%	6/17
Cleft palate/bifid uvula	29%	5/17
Dental abnormality	71%	12/17
Asymmetric hypomastia	29%	5/17
Heart defect	44%	7/16
Urogenital abnormality	29%	5/17
Sacral dimple	29%	5/17
Scoliosis	65%	11/17
Hip dysplasia	47%	8/17
Post-axial polydactyly	53%	9/17
Abdominal wall abnormality	12%	2/17
Anal atresia	53%	9/17
Neurology		
Seizures	24%	4/17
Hypotonia	47%	8/17
Brain abnormalities		
Dandy walker malformation (variant)	38%	5/13
Hypoplastic corpus callosum	62%	8/13
(Asymmetric) cerebellar hypoplasia	55%	6/11
(Asymmetric) enlarged ventricles	73%	8/11
Thin brain stem	30%	3/10
Abnormal gyration pattern frontal lobe	50%	5/10
Other		
Hearing loss	65%	11/17
(Blaschko) pigment abnormality	65%	11/17
Hypertrichosis	29%	5/17
Leg length discrepancy	41%	7/17
Malignancy	12%	2/17
Recurrent respiratory tract infections	53%	9/17
Thyroid hormone abnormality	35%	6/17

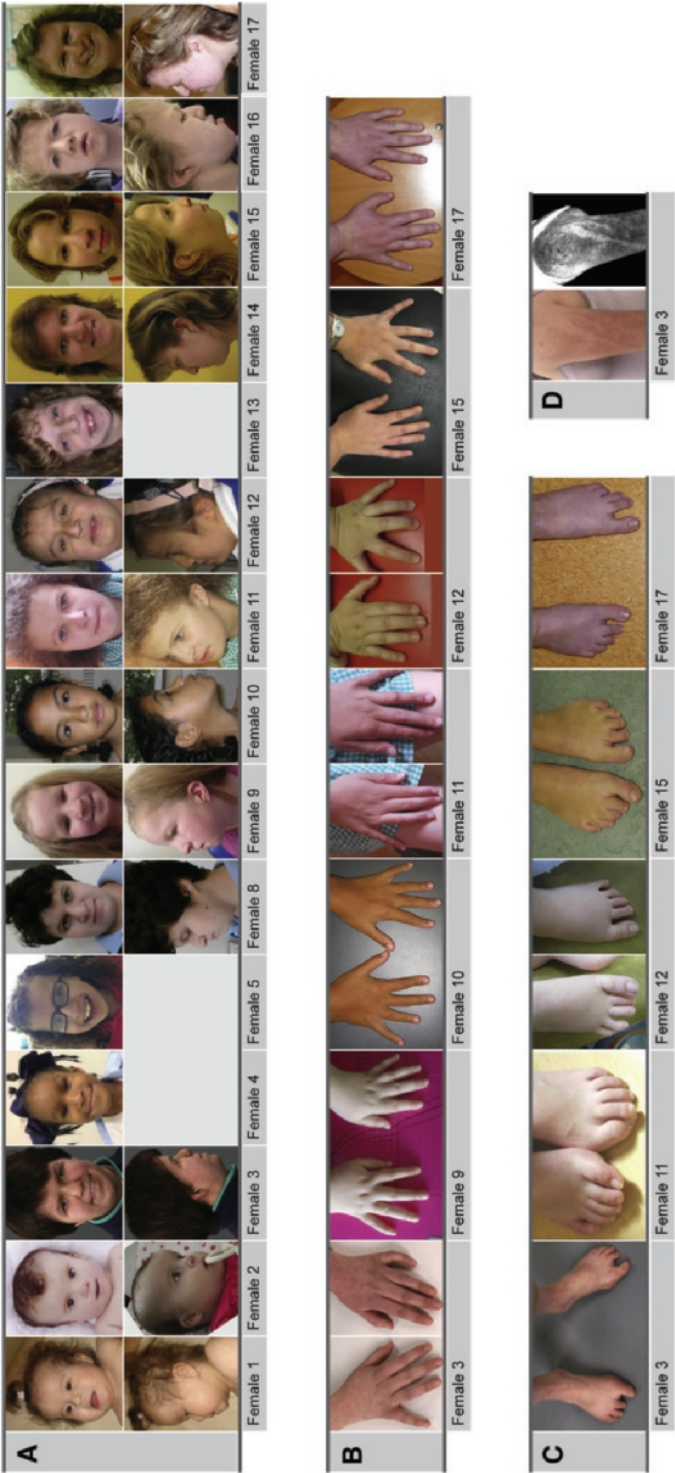


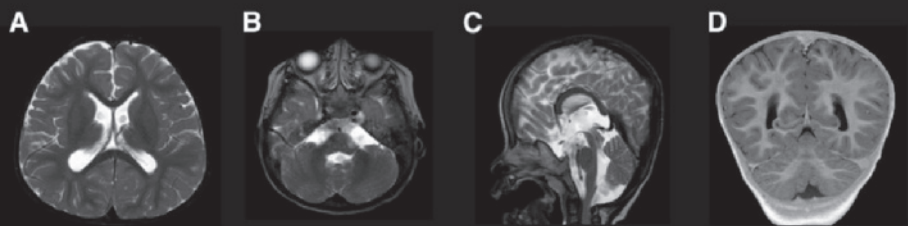
Figure 9.2: Clinical characteristics of females with de novo USP9X loss-of-function mutations. (A) Frontal and lateral photographs of females with de novo mutations in USP9X. Shared facial characteristics include facial asymmetry, prominent forehead, bitemporal narrowing, short palpebral fissures, low nasal bridge, prominent nose with flared alae nasi from adolescence age, thin upper lip, smooth and long philtrum, hanging full cheeks in early childhood, and low-set, posteriorly rotated, and dysplastic ears with attached lobule. (B) Photographs of the hands of seven affected females. Shared characteristics include ulnar deviation of 5th digit, tapered fingers, short 4th and 5th metacarpals, and post-axial polydactyly (simian crease present but not shown). (C) Photographs of the feet of five affected females. Shared characteristics include hallux valgus and sandal gap (pes cavus present but not shown). (D) Observed Blaschko lines of female 3, indicative for 11 of the affected females, suggestive of different X-inactivation pattern between tissues (functional mosaicism).

The X-linked *USP9X* encodes a structurally and functionally highly conserved deubiquitinating enzyme, containing a UBL (ubiquitin-like) and a catalytic ubiquitin specific protease (USP) domain.³¹²⁻³¹⁴ It is known to play an important role in neural development of both humans and mice and is required for fetal development.³¹⁵⁻³¹⁷ *USP9X* is highly expressed during embryogenesis and expression declines as cell fates become restricted.³¹⁶ The *USP9X* ortholog in *Drosophila*, fat facets (*faf*), has been shown to be important in cell polarity and cell fate of the developing eye in *Drosophila*.³¹⁸ A range of signaling proteins involved in different neurodevelopmental pathways including Notch, Wnt, TGF- β , and mTOR have been shown to interact with *USP9X*.^{313; 319-326} *USP9X* also has been described to act as both an oncogene and tumor-suppressor gene and is frequently found to be dysregulated in human cancer.^{313; 327; 328} Two of the affected females developed malignancy at a young age (22 and 8 years). Both acute lymphoblastic leukemia and osteosarcoma were treated successfully and have not reoccurred. To determine the risk and nature of particular malignancies in this new syndrome, further studies are required.

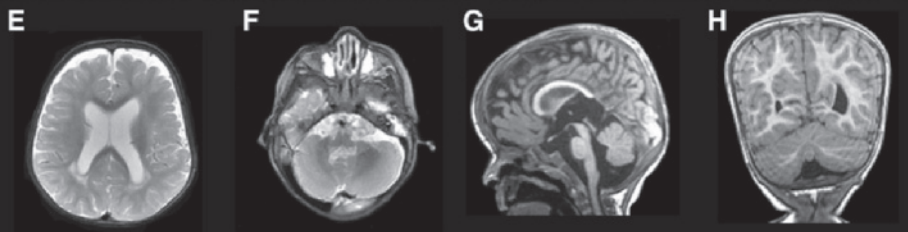
We observed pigment abnormalities along Blaschko lines and facial asymmetry, asymmetric abnormalities of the brain and breast, and asymmetric length of the legs (Figures 9.2A, 9.2D, and 9.3), all suggestive of a pattern of post-zygotic mosaicism or differential X-inactivation (XCI) between tissues (functional mosaicism).⁶² *USP9X* is one of the genes shown to escape XCI.^{329; 330} However, it is known that most of the genes that escape from XCI are not fully expressed from the inactivated X chromosome and instead show a partial escape.³³¹⁻³³³ Moreover, there is accumulating evidence for tissue-specific and developmental-stage-dependent differences in XCI and variability of escape of *USP9X*.^{332; 334-337} In the partial escaping genes, non-random XCI or skewing, as observed often in female carriers of an X-linked mutation, will only partially restore a normal phenotype.³³³ Consistent with this hypothesis, XCI was found to be skewed >90% in fibroblasts in three of the five of the tested females, but skewing was not related to disease severity (Supplemental Table 9.3). We note that a similar skewing pattern of XCI was observed recently in females with de novo mutations in *DDX3X* (MIM: 300160), another X-chromosomal gene that has been suggested to escape XCI and in which de novo mutations cause ID specifically in females.³³⁸

In one of the affected females, a predicted splice site mutation was identified. To evaluate whether this mutation indeed results in an aberrant transcript, we synthesized cDNA from RNA extracted from primary skin fibroblasts of both the affected female and a control. We amplified a fragment of 576 base pairs (bp) covering exon 20 to exon 22 by PCR. Electrophoretic separation showed two products of 576 and 455 bp in the sample from the affected female, and a single 576-bp product in the control. Sequencing of the smaller product revealed that this cDNA transcript from the affected female indeed lacked exon 21, confirming the truncating effect of the splice site mutation. Importantly, the level of the transcript was increased 4-fold when fibroblasts were treated with cycloheximide, strongly

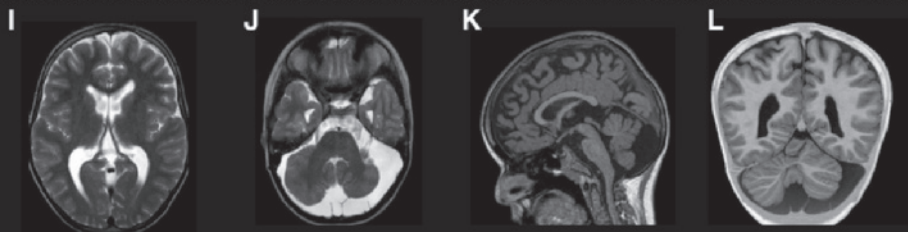
FEMALE 1



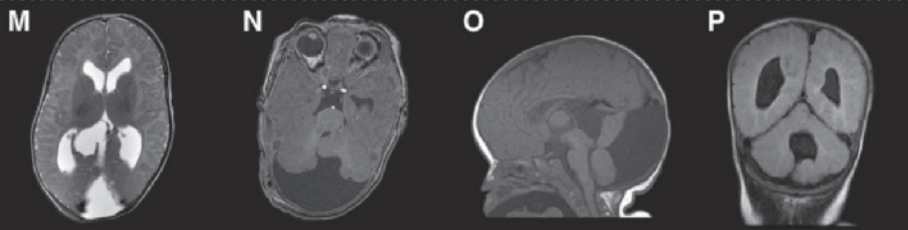
FEMALE 2



FEMALE 3



FEMALE 7



FEMALE 16

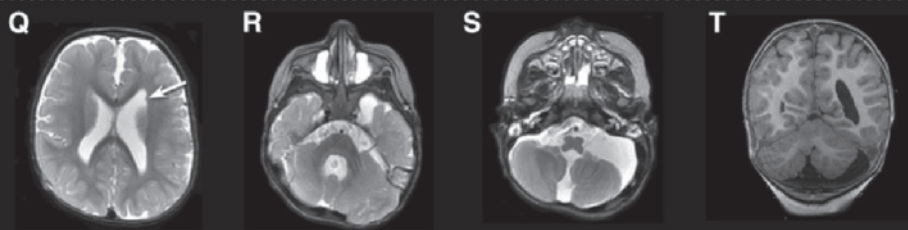


Figure 9.3: Representative MRI images from females 1, 2, 3, 7, and 16 with de novo *USP9X* loss-of-function mutations.

(A–D) Female 1 (2 years): MRI T2 axial (A, B) and sagittal (C) and T1 axial (D) sections show brachycephaly, mild enlargement of the lateral and 3rd ventricles; mild hypoplasia of cerebellar vermis and left cerebellar hemisphere; enlarged IV ventricle and cisterna magna with small retrocerebellar cyst; thin brain stem and mesencephalon; relatively small frontal lobes with somewhat simplified gyration; and short hypoplastic corpus callosum (both rostrum and splenium). (E–H) Female 2 (1.5 years): MRI T2 axial (E, F) and T1 sagittal (G) and coronal (H) sections show enlargement of the lateral ventricles, mild hypoplasia of cerebellar vermis and left cerebellar hemisphere; enlarged cisterna magna; thin corpus callosum, pons, mesencephalon, and brain stem; and broader and underdeveloped frontal gyri. (I–L) Female 3 (11 years): MRI T2 axial (I, J) and T1 sagittal (K) and axial (L) sections show asymmetric enlargement of the lateral ventricles; simplified convolutions of the frontal lobe gyri; hypoplasia of cerebellar vermis and left hemisphere; large cisterna magna and retrocerebellar cyst; and thin corpus callosum with hypoplasia of the rostrum. (M–P) Female 7: MRI T2 axial (M), T1 axial (N), T1 sagittal (O), and coronal FLAIR (P) sections show macrocephaly; enlargement of the lateral and 3rd ventricles with an interhemispheric cyst; dysplastic cerebellar hemispheres; dysplasia of the cerebellar vermis which is uplifted, with a high position of the tentorium and a large posterior fossa, typical of Dandy-Walker malformation; and thin and hypoplastic corpus callosum. (Q–T) Female 16 (2 years): MRI T2 axial (Q, R, S) and T1 coronal (T) sections show enlarged lateral ventricles; irregular gyri of the cerebral cortex with irregular depth of the sulci in frontal and perisylvian areas; small heterotopic nodule of gray matter (arrow) and thin and hypoplastic corpus callosum (both rostrum and splenium); hypoplasia of the anterior cerebellar vermis and left cerebellar hemisphere; enlarged cisterna magna and arachnoid cyst surrounding the cerebellum, especially at the left side; and mild hypoplasia of pons and brain stem. This female was identified with Sanger sequencing based on these brain abnormalities in combination with ID, dysmorphic features, and congenital abnormalities.

suggesting that the aberrant transcript was subjected to nonsense-mediated mRNA decay and as such leads to loss of function of this *USP9X* allele (Figure 9.1C). To study the effect of the heterozygous loss-of-function *USP9X* alleles on their mRNA expression and protein levels, we performed both qRT-PCR and immunoblot analysis of fibroblasts ($n = 4$) and lymphoblastoid cell lines (LCLs; $n = 1$) derived from affected females and both female and male controls (Figures 9.1D and Supplemental Figure 9.1). We found that expression of *USP9X* in affected females was reduced compared to control females in both fibroblasts and LCLs at both mRNA expression and protein levels. Although some cellular variability was evident, on average this decrease was significant ($p < 0.05$ by Student's *t* test) (Figures 9.1D and Supplemental Figure 9.1). There was no correlation between skewing of XCI and expression of mRNA and protein level. Whether the cells in affected tissue have benefit from the skewed XCI remains uncertain. The escape from XCI was supported by the fact that the average expression of *USP9X* mRNA in both control male fibroblasts and control male LCLs was ~50% of that in female controls. After quantification of protein levels in male control LCLs, similar levels were observed. The *USP9X* protein level in male control fibroblasts was increased to ~80% of that in female controls, but was still significantly less than the protein level in female controls. Intriguingly, these data thus reveal that affected females displayed reduced levels of *USP9X* compared to female controls but comparable levels to that in healthy control males. It will be important to expose whether these trends extend to other tissues, where the level of escape from XCI might not be comparable. Furthermore, characterization of different expression patterns between sexes (described for *USP9X* in brain thus far³³⁹) and/or that of protein levels of *USP9X* substrates will be important to ascertain as well.

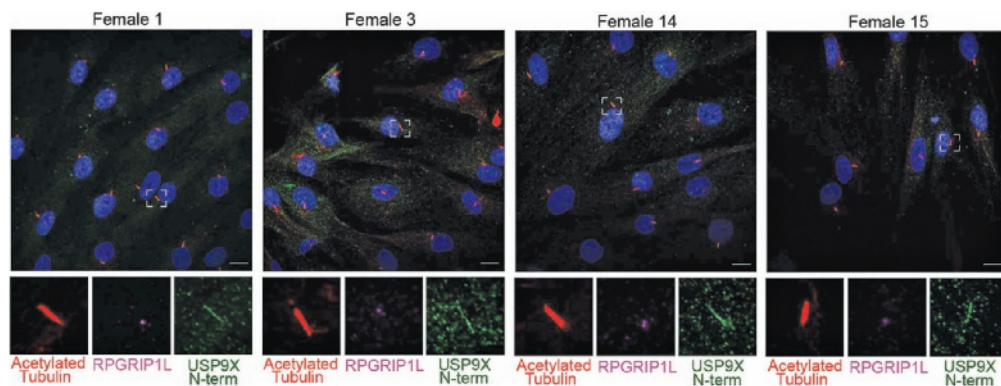


Figure 9.4: USP9X localizes to the primary cilium.

To induce ciliogenesis, control and affected female fibroblasts, matched for gender and age, were starved for 48 hr prior to immunofluorescence labeling. Endogenous USP9X is detected along the length of the axoneme of primary cilia, using an antibody against its N terminus (N-Term, shown in green) as compared to the ciliary markers RPGRIP1L (pink, denoting the ciliary transition zone at the base of the cilium) and acetylated α -tubulin (red, marking the ciliary axoneme). DAPI (blue) stains the nuclei; scale bars represent 10 μ m. Ciliated fibroblasts derived from affected females are shown here, USP9X localization in control fibroblasts is shown in Supplemental Figure 9.2.

In contrast to the severely disruptive *de novo* mutations in females, three milder mutations in *USP9X* have been reported in males without multiple congenital malformations. The mutations were transmitted by phenotypically normal females and resulted in ID, hypotonia, and behavioral problems in the males.³⁴⁰ An additional two missense mutations were identified after resequencing of *USP9X* in a cohort of 284 males with epilepsy.³⁴¹ For two *de novo* mutations reported in large autism cohorts, no specific gender information was described.³¹ The phenotype of the males differs notably from the observed phenotype in the affected females described here. They had ID and short stature, but lacked the multiple congenital malformations observed in affected females. Besides four missense mutations, one frameshift mutation has been reported in the males. This single frameshift mutation occurred within the last 50 nucleotides upstream of the last 3'-exon-exon junction, presumably escaping nonsense-mediated mRNA decay (NMD), and therefore results in a truncated protein lacking the last exon.^{342; 343} Because no truncating variants have been described in healthy controls in the ExAC database and no mutations causing loss of function of *USP9X* have been reported in males, we suspect that loss-of-function mutations could be lethal in males. This hypothesis is further supported by the fact that the absence of *Usp9X* in male mice is embryonically lethal.³⁴⁴ In contrast, all but one of the affected females we report here have protein-truncating mutations and deletions. In one female, we identified a *de novo* missense mutation, located in the catalytic domain of the protein. Given the fact that this female was phenotypically comparable with the other females, it is likely that this specific missense mutation leads to loss of function of the protein. We hypothesize that in addition to complete loss-of-function mutations, such as protein-truncating mutations and deletions, a small subset of specific missense mutations will also lead to disease in females.

Affected females presented with symptoms that overlap with CHARGE syndrome (MIM: 214800) (*CHD7* [MIM: 608892] tested in four of the females) and with the clinical spectrum of some known ciliopathy syndromes, such as Bardet-Biedl, Meckel-Gruber, and Joubert syndromes.³⁴⁵ Therefore, we investigated whether heterozygous protein-truncating mutations result in the disruption of ciliary structure, formation, or trafficking in fibroblasts of four of the affected females we had available (females 1, 3, 14, and 15). First, we determined subcellular localization of endogenous USP9X in both controls and affected-female-derived fibroblasts under serum starvation to induce ciliogenesis, as well as in serum-rich conditions.^{346; 347} USP9X showed diffuse cytoplasmic staining with areas enriched with puncta consistent with its described association with protein and vesicle trafficking.³⁴⁸ Importantly, upon the induction of ciliogenesis in the fibroblasts, USP9X was indeed found to localize to the cilium. This ciliary localization was observed along the length of the ciliary axoneme of most fibroblasts, and comparable in cells from affected females and from age- and gender-matched controls (Figure 9.4). This localization was significantly decreased with siRNA knockdown of *USP9X* indicating specificity of the signal (Supplemental Figure 9.2). We were unable to observe any significant differences in ciliogenesis, ciliary length, or ciliary trafficking between fibroblasts from affected females when compared to controls (Supplemental Figure 9.3), and siRNA knockdown of *USP9X* did not impair ciliogenesis in fibroblasts. This suggests that USP9X dosage is not critical to the generation of primary cilia in these fibroblasts, despite localization of USP9X in their cilia. It is therefore more likely that USP9X-regulated signal transduction pathways mediated by the primary cilium are more subtly disturbed, and/or that this disturbance is spatiotemporally restricted to the tissues affected in this specific phenotype, possibly due to tissue-specific and developmental-stage-dependent differences in XCI and variability of escape of USP9X. Future studies utilizing dedicated cell-based or animal models will be necessary to evaluate these mechanisms.

In conclusion, this study defines a recognizable X-linked ID/DD syndrome with associated multiple congenital malformations and brain abnormalities specific to females, caused by de novo loss-of-function mutations in *USP9X*, a gene known to escape X-inactivation. The phenotypic characteristics overlap with ciliopathy conditions and USP9X localization along the length of the ciliary axoneme of fibroblasts indicates a role in de-ubiquitination of ciliary proteins, which could contribute to the disease pathogenesis of this specific syndrome.



10

***De novo* loss-of function mutations in WAC cause a recognizable intellectual disability syndrome and are associated with learning deficits in Drosophila**

This chapter has been published as:

D. Lugtenberg*, **M.R.F. Reijnders***, M. Fenckova*, E.K. Bijlsma, R. Bernier, B. W.M. van Bon, E. Smeets, A.T. Vulto-van Silfhout, D. Bosch, E.E. Eichler, H.C. Mefford, G.L. Carvill, E.M.H.F. Bongers, J.H.M. Schuurs-Hoeijmakers, C.A. Ruivenkamp, G.W.E. Santen, A.M.J.M. van den Maagdenberg, C.M.P.C.D. Peeters-Scholte, S. Kuenen, P. Verstreken, R. Pfundt, H.G. Yntema, P.F. de Vries, J.A. Veltman, A. Hoischen, C. Gillissen, B.B.A. de Vries, A. Schenck*, T. Kleefstra*, L.E.L.M. Vissers*

European Journal of Human Genetics (2016) 24, 1145-53

* These authors contributed equally

Abstract

Recently *WAC* was reported as a candidate gene for intellectual disability (ID) based on the identification of a *de novo* mutation in an individual with severe ID. *WAC* regulates transcription-coupled histone H2B ubiquitination and has previously been implicated in the 10p12p11 contiguous gene deletion syndrome. In this study, we report on 10 individuals with *de novo* *WAC* mutations which we identified through routine (diagnostic) exome sequencing and targeted resequencing of *WAC* in 2326 individuals with unexplained ID. All but one mutation was expected to lead to a loss-of-function of *WAC*. Clinical evaluation of all individuals revealed phenotypic overlap for mild ID, hypotonia, behavioral problems and distinctive facial dysmorphisms, including a square-shaped face, deep set eyes, long palpebral fissures, and a broad mouth and chin. These clinical features were also previously reported in individuals with 10p12p11 microdeletion syndrome. To investigate the role of *WAC* in ID, we studied the importance of the *Drosophila* *WAC* orthologue (*CG8949*) in habituation, a non-associative learning paradigm. Neuronal knockdown of *Drosophila* *CG8949* resulted in impaired learning, suggesting that *WAC* is required in neurons for normal cognitive performance. In conclusion, we defined a clinically recognizable ID syndrome, caused by *de novo* loss-of-function mutations in *WAC*. Independent functional evidence in *Drosophila* further supported the role of *WAC* in ID. On the basis of our data *WAC* can be added to the list of ID genes with a role in transcription regulation through histone modification.

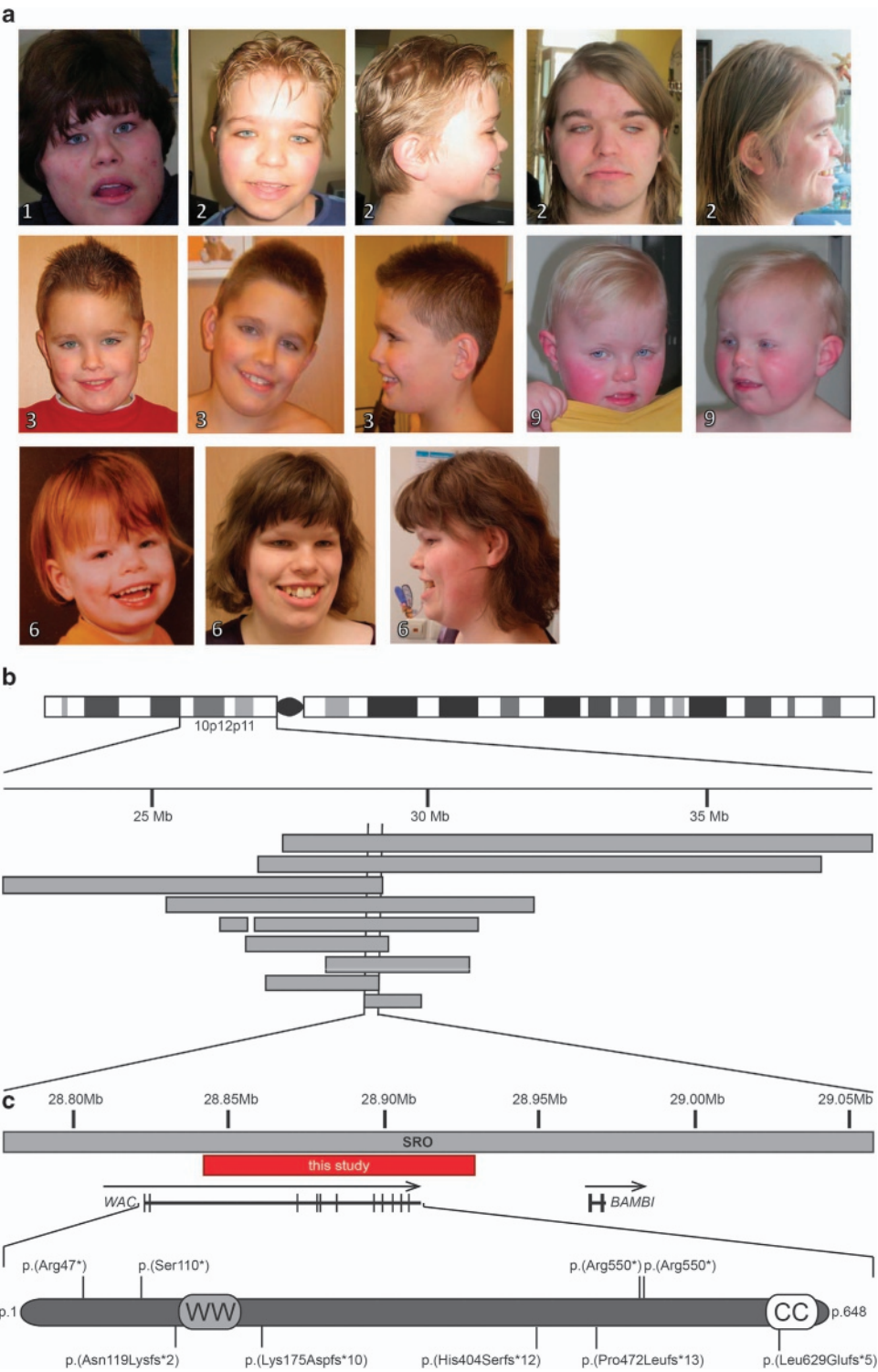
Introduction

Intellectual disability (ID) is a heterogeneous disorder, both clinically and genetically. To date, 4650 genes have been associated with ID and novel genes are still being identified. The introduction of triobased whole-exome sequencing (WES) in individuals with ID has proven to be a valuable approach for the identification of novel ID genes, especially for those individuals who do not show a clinical recognizable syndrome.^{28; 29} In addition to the identification of mutations in known disease genes, WES has facilitated the identification of candidate ID genes. To establish the pathogenicity of mutations in such candidate ID genes, it is essential to identify additional individuals with an overlapping phenotype and a mutation in the same gene.³⁴⁹⁻³⁵¹ With increasing availability of WES in routine diagnostics⁶ as well as technological advances facilitating targeted resequencing of candidate ID genes in larger cohorts of samples,³⁵² chances of finding such additional individuals are increasing. Furthermore, supporting evidence and insights into underlying mechanisms can be obtained from functional studies in cell or animal models.^{353; 354} Previously, we and others separately reported on an individual with a *de novo* mutation in the 'WW domain-containing adapter with coiled-coil' (WAC) gene using trio-based exome sequencing.^{28; 355} The mutations were reported as potential cause of disease, based on mutation severity, protein function, its expression in fetal stages and high expression in adult brain.^{28; 355; 356} WAC encodes a protein-regulating transcription-coupled histone H2B ubiquitination and contains two evolutionary conserved domains, including an N-terminal WW domain interacting with RNA polymerase II and a C-terminal coiled-coil domain promoting the RNF20/RNF40's E3 ligase activity for ubiquitination at active transcription sites.^{357; 358} Furthermore, the RNF20/40/WAC complex may have a role in cell cycle checkpoint activation upon genotoxic stress.³⁵⁷ In addition, WAC has previously been implicated in ID based on the finding that deletions of chromosome 10p12p11 result in a contiguous gene deletion syndrome, for which the shortest deleted region contains two genes, WAC and *BAMBI*. All individuals with a deletion of at least these two genes were reported to have a similar phenotype including ID, behavior problems and dysmorphic features, supporting a disease cause of WAC heterozygous loss-offunction.³⁵⁹⁻³⁶² Although it may well be hypothesized that WAC haploinsufficiency may explain the ID phenotype observed in the 10p12p11 contiguous gene syndrome, and ID in individuals with mutations in this gene, detailed evidence to support this hypothesis is lacking. In the present study we aimed to identify additional individuals with *de novo* mutations in WAC by using different sequencing strategies to define the clinical spectrum associated with WAC haploinsufficiency. Finally, to address the role of WAC in cognition, we investigated the role of the *Drosophila* WAC orthologue in habituation, a form of nonassociative learning.

Results

Identification of individuals with de novo CNVs affecting WAC

After the identification of the *de novo* mutation c.139C>T (NM_016628.3) leading to nonsense mutation p.(Arg47*) in Individual 1 (as reported before²⁸), we set out to find additional



◀ **Figure 10.1: Individuals with WAC mutations.**

(A) Frontal and lateral photographs of individuals with *de novo* mutations in WAC. All individuals shared overlapping facial dysmorphisms including a square-shaped face, long palpebral fissures, broad mouth and broad chin. Additional features included deep set eyes, epicanthal folds and short philtrum in individual 1 (photograph at the age of 19 years); low posterior hairline, broad forehead, simple ears, hypertelorism, deep set eyes, low-set full eyebrows, synophrys, deep nasal bridge, flat nose, bifid tongue and broad gums in individual 2 (photographs at the age of 12 years and 23 years); brachycephaly, posterior ear creases, broad forehead, prominent antihelix, low-set full eyebrows, synophrys and prominent teeth in individual 3 (photographs at the age of 4 years and 9 years); prominent antihelix, frontal bossing, dental crowding, broad teeth and high palate in individual 6 (photographs at the age of 3 years and 20 years) and prominent antihelix and deep set eyes in individual 9 (photographs at the age of 3 years). Photographs were published with consent. (B) The genomic region involved in the 10p12p11 contiguous gene deletion region with the previously published microdeletions, represented by gray horizontal bars.^{359–361} (C) Detailed view of the smallest region of overlap (SRO) and the deletion described in this study, represented by a red horizontal bar. In addition, *de novo* mutations in WAC (NM_016628.3) reported in this study are shown according to their relative position at protein level.

individuals with mutations affecting WAC to obtain more evidence for its involvement in ID. Systematic analysis of DECIPHER and ECARUCA, two databases collecting clinically relevant copy number variants (CNVs), for CNVs affecting WAC yielded one small deletion (Individual 2; Figure 10.1). This deletion, hg19 chr10:g.(?_288422777)_(28929097_?)del, disrupted the coding sequence of WAC by deletion of exons 5–14 (NM_016628.3). The shortest region of deletion overlap of the chromosome 10p12p11 contiguous gene deletion syndrome was previously determined by nine deletions ranging in size between 0.99 and 10.66 Mb.^{359–362} Comparison of the deletion in Individual 2 to the shortest region of deletion overlap indicates WAC as a only remaining candidate gene for the ID phenotype (Figure 10.1B).

Diagnostic exome sequencing in individuals with neurodevelopmental disorders to identify de novo point mutations in WAC

In routine diagnostic trio-based exome sequencing for individuals with unexplained ID, performed as described before,²⁸ we identified four additional individuals with *de novo* loss-of-function mutations in WAC (NM_016628.3): c.329C>A, p.(Ser110*) in Individual 3; c.1885_1886del, p.(Leu629fs) in Individual 4; c.356dup, p.(Asn119fs) in Individual 5 and c.1648C>T, p.(Arg550*) in Individual 6 (Figure 10.1C). In addition, two more individuals with *de novo* mutations were identified by trio-based exome sequencing of a large cohort of individuals with autism spectrum disorder:^{31; 114} c.523_524del, p.(Lys175fs) in Individual 7 and c.1209_1212del, p.(His404fs) in Individual 8. Whereas exome sequencing had identified a second *de novo* mutation in Individual 1 (MIB1; NM_020774.2:c.521G>A; p.(Arg174His)),²⁸ no further clinically relevant *de novo* mutations were identified in Individuals 3–8, leaving WAC haploinsufficiency as the most likely candidate to explain disease.

Targeted resequencing of WAC in an ID cohort identifies additional de novo mutations

On the identification of multiple *de novo* loss-of-function mutations in WAC, we performed targeted resequencing of this gene in a cohort of over 2300 individuals with unexplained ID

using MIPs as described before.³⁵² This cohort was selected from the in-house collection of the Department of Human Genetics of Radboud University Medical Center containing individuals with unexplained ID. All candidate loss-of-function mutations as well as highly conserved missense mutations (PhyloP>5) were validated by standard Sanger sequencing approaches. For assessing the *de novo* occurrence of validated mutations, DNA from the parents was tested. This targeted screen identified two additional *de novo* truncating mutations: c.1415del, p.(Pro472fs) in Individual 9 and c.1648C>T, p.(Arg550*) in Individual 10 (Figure 10.1C).

WAC mutation spectrum in control individuals

Of all ten *de novo* mutations identified, nine are predicted to directly result in nonsense-mediated decay of the RNA transcripts; the *de novo* frameshift in Individual 4 is located in the last exon, suggesting it may skip nonsense-mediated RNA decay. None of the *de novo* mutations are reported in our in-house variant database containing Exome sequencing variants detected in 5031 individuals, nor in ExAC, a large database collecting NGS variants in over 60 000 exomes as proxy for variant allele frequencies in the general population.³⁶³ The latter, however, does contain three other, presumable loss-of-function, variants (by insertion–deletion events), each observed only once in ~100 000 alleles. Whereas these three variants have not been validated by Sanger sequencing, thereby possible being sequencing artefacts rather than true mutations, this observation may also reflect the very mild end of the ID spectrum in the general population.

Clinical spectrum associated with WAC haploinsufficiency

Clinical evaluation of all individuals with *de novo* loss-of-function mutations in *WAC* showed distinct phenotypic overlap (Table 10.1; Supplemental Information (clinical descriptions); and Supplementary Table 1.1). All, but one individual, had ID. The range of ID observed ranged from mild-to-severe and was accompanied by language and motor delay. In addition, individuals showed a variety of neurological problems including hypotonia (6/9), with remarkable manifestation in the oral region resulting in dysarthria, and behavioral problems (10/10). The latter recurrently included autism (4/9), anxiety (3/10), concentration disorder (4/10) and/or sleep disturbance (6/10). Other overlapping features consisted of unexplained reduced vision (3/9) and respiratory problems (7/9) with recurrent respiratory infections reported most often (5/7). Notably, all individuals had overlapping facial dysmorphisms consisting of a square shape of the face, deep set eyes, long palpebral fissures, broad mouth and broad chin (Figure 10.1A).

Panneuronal knockdown of the Drosophila WAC orthologue results in learning deficit

To obtain independent evidence for the involvement of *WAC* in the ID phenotype of the described individuals, we decided to study the functional consequences of *WAC* knockdown

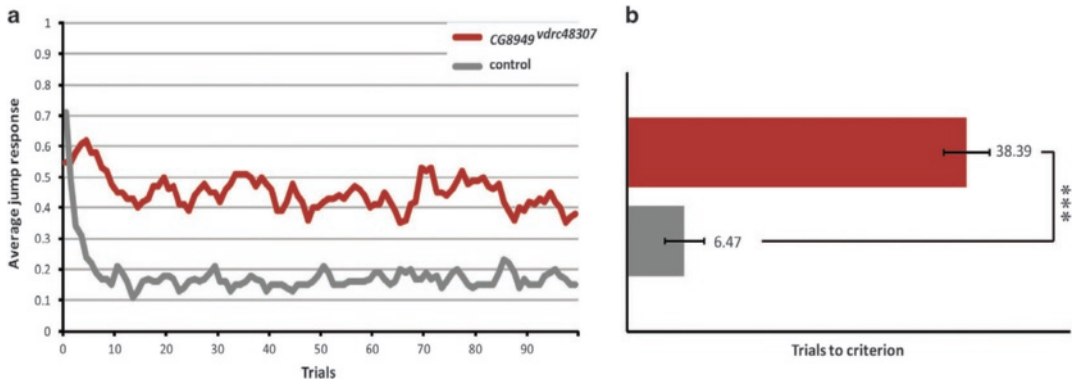


Figure 10.2: Knockdown of the *Drosophila* WAC orthologue CG8949 affects non-associative learning in the light-off jump reflex habituation paradigm.

Jump responses of 3-to-7-day-old individual male flies were induced by repeated light-off pulses (100 trials) with a 1s inter-trial interval. CG8949 knockdown flies (CG8949^{vdr48307}; genotype: 2xGMR-wlR/+; UAS-CG8949^{vdr48307}/elav-Gal4, UAS-Dicer-2) are plotted in red and genetic background control flies are plotted in dark grey. Habituation was scored as the mean number of trials required to reach the non-jump criterion (Trials To Criterion, TTC). Main effects of genotype (mutant vs control), day and test system on log-transformed TTC values were tested using linear model regression analysis.³⁶⁰ (A) Average jump response across 100 light-off trials. (B) Mean TTC of CG8949^{vdr48307} (TTC=38.39, n=54) vs mean TTC of control flies (TTC=6.47, n=49). Quantification of average jump responses revealed that flies with pan-neuronally induced CG8949 knockdown habituated significantly slower (***) $P < 0.001$, linear model regression analysis).

using *Drosophila* as a model. The *Drosophila* genome contains a previously uncharacterized WAC orthologue named CG8949, not be confused with the unrelated *Drosophila* WAC (*wec Augmin*) gene. CG8949 codes for several protein isoforms, the longest one consisting of 876 amino acids, and shows the highest expression in adult ovaries and the larval central nervous system.^{364; 365} WAC is 26% identical and 39% conserved over the C-terminal 588 amino acids of the fly protein, with sequence similarity distributed over the whole protein, further characterized by clusters of short sequences of high conservation for the important functional motifs of the protein.³⁵⁸ We investigated the role of the *Drosophila* WAC orthologue in lightoff jump reflex habituation paradigm. Habituation is a simple form of non-associative learning, in which an initial strong behavioral response towards a repeated, non-threatening stimulus gradually wanes. It provides a filtering mechanism, which is an important prerequisite for higher cognitive functioning.^{366; 367} Using the light-off jump reflex habituation, we have previously identified learning deficits in number of *Drosophila* ID models.^{354; 368; 369} Two independent inducible RNAi lines targeting the *Drosophila* WAC orthologue CG8949 (*vdr48307* and *vdr107328*) and their corresponding genetic background control lines (*vdr60000* and *vdr60100*) were obtained (Vienna *Drosophila* RNAi Center³⁷⁰) and fly stocks were kept under standard conditions. Expression of CG8949 was specifically downregulated in neurons using the UAS-Gal4 system. The efficiency of ubiquitous RNAi knockdown was measured using qPCR on RNA isolated from *Drosophila* brains of third instar larvae, representing the tissue and developmental stage with the highest expression of CG8949

Table 10.1 Clinical details of individuals with WAC mutations

	Individual 1	Individual 2	Individual 3	Individual 4	Individual 5	Individual 6	Individual 7	Individual 8	Individual 9	Individual 10
Gender	Female	Male	Male	Male	Female	Female	Female	Female	Female	Male
Age at last visit	22 years	12.5 years	9 years	17 months	3 years	20 years	6 years	9 years	12 years	7 years
Mutation WAC (NM_016628.3)										
cDNA change	c.139C>T	partial gene	c.329C>A	c.1885_1886del	c.356dup	c.1648C>T	c.523_524del	c.1209_1212del	c.1415del	c.1648C>T
Amino acid change	p.(Arg47*)	deletion	p.(Ser110*)	p.(Leu629fs)	p.(Asn119fs)	p.(Arg550*)	p.(Lys175fs)	p.(His404fs)	p.(Pro472fs)	p.(Arg550*)
Chromosome position	Chr10:28824551	Chr10:g. (?_288422777)_ 28829097_?del	Chr10:28872382	Chr10:28908476_ 28908477	Chr10:28872409	Chr10:28905193	Chr10:28879674_ 28879675	Chr10:28899671_ 28899674	Chr10:28900827	Chr10:28905193
Growth										
Birth weight (g)	NR	3600 (>+2 SD)	3600 (0 SD)	NR	2920 (0 SD)	3070 (-1 SD)	NR	NR	3380 (0 SD)	4500 (+2.5 SD)
Height (cm)	NR	149.7 (-1.1 SD)	131.3 (-1.5 SD)	75.4 (+2 SD)	87 (-3 SD)	156.5 (-2.5 SD)	+0.86 SD	+ 2.88 SD	161.0 (+0.75 SD)	114.8 (-1.75 SD)
Weight (kg)	NR	48.8 (+2.1 SD)	31.5 (+0.75 SD)	10.0 (+2 SD)	13 (+0.7 SD)	75.5 (+3.5 SD)	+1.7 SD	-2.9 SD	61.2 (+ 2 SD)	26.2 (+3 SD)
Head circumference (cm)	NR	56.3 (+1.1SD)	52.0 (-0.5 SD)	48.5 (0 SD)	50.4 (-0.6 SD)	54 (-1 SD)	+ 2.9 SD	+ 0.53 SD	56.0 (+ 1.5 SD)	54.9 (+2 SD)
Development										
Intellectual disability	Severe	Mild (IQ 58)	Mild (IQ 65)	Not assessed owing to young age	Mild	Moderate (IQ 44)	No (IQ 89)	Mild (IQ 61)	Mild (IQ 65)	Mild (IQ 61)
Adaptive ability										
Speech delay	+ No speech	+ Dysarthria	+ 2.5 years: first words Dysarthria	+ 1.5 years: no words, only babbling	+ Dysarthria	+ 2 years: first words Dysarthria	- 20 months: first words	+ 5 years: first words	+ 5 years: first words	+ 1.5 years: first words
Motor delay	+ 34 months: first steps	+ 24 months: first steps	+ 24 months: first steps	+ 1.5 years: walks with support	+ 36 months: walks unsupported	+ 2 years: first steps	+ 20 months: first words	NR	+ 22 months: first steps	+ 27 months: first steps
Behaviour										
Anxiety	-	+	-	-	+	+	-	-	+	-
Concentration	-	-	+	-	+	+	-	-	-	+
problems/hyperactivity										
Sleep disturbances	+	+	-	+	+	+	-	-	+	-
Autism	-	+	-	-	NR	+	+	+	-	-
Other	Self-mutilation	-	-	Age 1 year: lack of eye contact	-	-	-	-	-	-
Neurological										
Epilepsy	+	-	-	-	-	-	-	-	-	-
Oral hypotonia	NR	+	+	+	+	+	-	-	+	+
Hypotonia	NR	+	+	+	+	-	-	-	+	+
Toe walking	+	-	+	-	NR	-	NR	NR	-	+

(Continues on the next page)

Ocular									
Vision	Normal	Normal	Unexplained reduced vision	To further establish: lack of eye contact and staring	Hypermetropia (+3.75 Dpt)	Unexplained reduced vision	Normal	Hyper/hypometropia	CVI
Strabismus	-	-	+	-	-	+	-	-	+
Pulmonal									
Abnormal breathing pattern	+	-	-	-	NR	-	NR	±a	-
Recurrent respiratory infections	NR	+	+	+	-	+	NR	±a	+
Asthma	-	+	+	-	-	-	NR	±a	-
Extremities									
Brachydactyly fingers	NR	-	+	-	-	+	NR	NR	+
Padres plano valgi	NR	-	+	+	-	+	NR	NR	+
Fetal finger pads	NR	-	+	-	-	-	NR	NR	+
Other	-	Broad first digits	Sandal gap	-	Short hands and feet	Tapering fingers	-	-	Hip dysplasia
			Simian crease						Simian crease
Other									
Neonatal feeding difficulties	+	-	+	-	+	-	-	-	+
Other clinical features	Hernia diaphragmatica, pubertas tarda	Laryngomalacy	Fever convulsions, fatigue, phinosis, hypertrichosis	Large frontal fontanel	Low IgG class 2 and 3	Fatigue, paroxysmal atonia	-	Kidney problems	Inverted nipples
			back						
Additional investigations									
MRI/CT	Normal	Normal	Megacisterna magna	Normal	Normal	Normal	NR	Abnormal (no details available)	Mild asymmetrical right frontotemporal hemisphere
Muscle biopsy	-	-	Low mitochondrial activity	-	-	Normal	NR	NR	-
Genetic analysis	250k SNP array: - ; MECP2: - ; RAI1: - ; TCF4: - ; Angelman methylation: -	FMR1: - ; RAI1: -	Subtelomeric regions: - ; 250k SNP array: - ; OPAT: - ; POLG: - ; RYR1: -	Cytoscan HD array: 9p22.1 gain & Xp22.31 gain, maternally inherited;	FMR1: - ; RAI1: - ; EMT1: - ; Cytoscan: 6Mb region of homozygosity chromosome 6q14.2q15	Angelman methylation - ; FMR1: - ; CACNA1A: - ; array CGH: -	NR	NR	FMR1: - ; Subtelomeric regions: -
Other (clinically relevant) de novo variant(s) by exome sequencing	MIB1(NM_020774.2) c.521G>A; p.(Arg174His)	not tested	none	none	none	none	none	none	not tested

Abbreviations: CVI, cerebral visual impairment; Dpt, diopter; IQ, intelligence quotient; NR, not reported. a ± indicates that Individual 8 is reported to have pulmonal problems, but this is not further specified.

allowing for an efficient detection of expression differences. There was no significant *CG8949^{vdrC107328}*-mediated knockdown on *CG8949* expression levels (88% remaining gene expression; $P=0.16$, student's *t*-test; Supplemental Table 10.2). In contrast, *CG8949^{vdrC48307}*-mediated RNAi reduced levels of *CG8949* to 58% remaining gene expression ($P<0.01$, Student's *t*-test; Supplemental Table 10.2).

Flies were exposed to series of 100 short (15 ms) light-off stimuli with 1 s interval between stimuli. Both control and *CG8949^{vdrC48307}* knock-down flies showed good initial jump response; there was no significant difference between the initial startle response of *CG8949^{vdrC48307}* and control flies (*t*-test, $P = 0.469$). Whereas, control flies quickly habituated to the repeated light-off stimuli, *CG8949^{vdrC48307}* knock-down animals failed to adapt their behavioral response and retained high average jump response throughout the whole experiment (Figure 10.2). This defect was statistically significant (fold-change = 5.93; $P = 5.21 \times 10^{-12}$). No habituation defects were seen in the *CG8949^{vdrC107328}* knock-down flies (data not shown), as was to be expected based on the insufficient mRNA knockdown.

Discussion

Here, we report the identification of a novel clinically recognizable syndrome caused by haploinsufficiency of *WAC*. All but one patient showed mild ID, with speech and motor delay, whereas one had an overall more severe ID phenotype, epilepsy and an absence, rather than delay, of speech. All patients had neurological problems including hypotonia and a variety of behavioral problems including autism, anxiety, concentration problems, sleep disturbance and/or self-mutilation. Notably, all patients had overlapping facial dysmorphisms consisting of square shape of the face, deep set eyes, long palpebral fissures, broad mouth and broad chin. Complementary experimental evidence in *Drosophila* showed a role of the evolutionarily conserved *WAC* proteins in cognitive processes and a role for the *Drosophila* *WAC* orthologue in non-associative learning.

Previously, another individual was reported with a truncating mutation in *WAC*³⁵⁵ who shows a phenotype similar to the individuals reported in this study (Supplemental Table 10.1). Also, a large-scale study aiming at the identification of genetic causes underlying developmental disorders recently reported the identification of one *de novo* nonsense mutation in *WAC* but further clinical details of this individual were lacking, hampering detailed phenotypic comparison.³² Interestingly, three of our patients were negative tested for *RAI1*, known to cause Smith–Magenis syndrome.³⁷¹ The coarse facial appearance as well as the ID with variable behavior problems of the individuals with *WAC* mutations have similarities with individuals reported with Smith–Magenis syndrome.³⁷²

Interstitial deletions including *WAC* were previously described and associated with ID.^{359–362} Wentzel et al.³⁵⁹ presented six individuals with an interstitial deletion at 10p12p11, all sharing a region of overlap including two genes: *BAMBI* and *WAC*. All individuals were reported to

have developmental delay, abnormal behavior and facial dysmorphic features including a bulbous nasal tip, deep set eyes, synophrys/thick eyebrows and full cheeks. This phenotype is highly similar to the phenotype observed in the current individuals and consistent with our finding that loss of *WAC* causes ID and the characteristic facial dysmorphism in 10p12p11 microdeletion syndrome (Supplemental Table 10.1). In this 10p12p11 microdeletion syndrome, cardiac abnormalities have frequently been reported (7/9 individuals) and heterozygosity of two other genes, *LYZL1* and *SVIL*, has been suggested to contribute to the development of these cardiac abnormalities.³⁵⁹ This is in line with the fact that in none of our individuals cardiac abnormalities were present. Epilepsy has been reported in two out of nine individuals with deletion of 10p12p11 and is present in only one of our individuals (Individual 1). This more severely affected individual carried also a *de novo* *MIB1* mutation. This variant has been reported twice in ExAC, containing NGS variants in healthy controls of several ethnicities.³⁶³ Moreover, one missense and one nonsense mutations in *MIB1* were identified previously and segregated each in two large dominant families with affected individuals with cardiomyopathy, but without ID.³⁷³ Therefore, a contribution of the second mutation in *MIB1* as potential modifier of the more severe phenotype is unlikely. The more severe phenotype may be caused by other yet unknown potential genetic modifier(s) or reflects the severe end of the clinical spectrum caused by *WAC* haploinsufficiency.

Human *WAC* encodes a protein containing a WW domaincontaining adapter and a coiled-coil region. *WAC* is an evolutionary conserved protein, but its exact function is still unknown. Protein–protein interaction studies suggest a role in the regulation of histone H2B ubiquitination and gene transcription.³⁵⁷ Interaction of *WAC* via the coiled-coil domains with RNF20 and RNF40 activates UBE2A-mediated H2B ubiquitination. On the basis of this interaction, RNF20 and RNF40 could be considered as candidate genes for ID. Interestingly, *de novo* mutations in both genes have been described in individuals with autism and unrelated unaffected siblings.³¹ A clear relation to a clinical phenotype, if any, remains to be established. RNA polymerase II recruits *WAC* to active transcription sites by binding to the WW domain. On the basis of sequence homology it has also been suggested that *WAC* is involved in RNA processing or transcription.³⁵⁸ Besides co-localization of *WAC* with splicing factor SC35 in nuclear speckles, there is, however, currently no further evidence for a role in RNA splicing. The localization of *WAC* in the nucleus supports its suggested function in gene transcription via UBE2A histone H2B ubiquitination. Although *WAC* and UBE2A lack a direct interaction, they are both likely to function within a protein complex important for histone H2B ubiquitination and transcription regulation.³⁵⁷ Interestingly, for *UBE2A*, microdeletions and point mutations have been associated in males with an X-linked inherited clinical syndrome characterized by ID, seizures, absent speech, urogenital and skin anomalies. Recent work on *Drosophila* has uncovered a novel role of UBE2A in clearance of defective mitochondria from the synaptic compartment and in synaptic plasticity.³⁷⁴ On the basis of the suggested interaction of *WAC* and UBE2A being part of the same complex, and their

clinical phenotypes both including ID, we experimentally addressed whether CG8949, like fly UBE2A, is required for synaptic vesicle cycling, mitochondrial functioning and morphology.³⁷⁴ We found all these processes unperturbed in our *Drosophila* model (Supplemental Figure 10.1). Despite the unavailability of the second CG8949 knockdown fly as an independent confirmation of the habituation phenotype, our data supports the role of WAC in cognition by the deficit in nonassociative learning.

In summary, we describe a clinically recognizable syndrome owing to loss-of-function mutations in WAC, which is characterized by mild ID, hypotonia, behavioral problems and facial dysmorphisms consisting of square shape of the face, deep set eyes, long palpebral fissures, broad mouth and broad chin. Complementary experimental evidence in *Drosophila* showed a role of the evolutionarily conserved WAC proteins in cognitive processes and a role for the *Drosophila* WAC orthologue in nonassociative learning. WAC is a component of the evolutionary conserved histone H2B ubiquitination complex, regulating gene transcription. On the basis of our results, WAC can be added to the growing list of genes involved the H2B ubiquitination complex leading to cognitive defects.

Material and methods

Diagnostic exome sequencing

Individuals 1, 3, 4, 5 and 6 were ascertained through family-based WES in a diagnostic setting using techniques as described before.²⁸ All clinically relevant candidate *de novo* mutations were validated using Sanger sequencing, and subsequently tested for absence in parental DNA samples. Individual 1 was previously reported as part of a large study assessing the clinical utility of WES, in which she was also identified to have a second *de novo* mutation in *MIB1*.²⁸ Individuals 7 and 8 were identified in a large multicenter study to establish the contribution of *de novo* coding mutations to autism spectrum disorder.^{31;}
¹¹⁴ Apart from the *de novo* mutation in WAC these two patients were not reported to have additional *de novo* mutations.³¹

Database searches for copy number variations disrupting WAC

We systematically searched for individuals with small deletions including WAC in our in-house database and international databases such as the database of the European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA) and the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER).^{375; 376}

Targeted resequencing of WAC

Upon identification of the *de novo* mutation in Individual 1, targeted resequencing was performed on a cohort of 2326 patients with unexplained ID using molecular inversion probes (MIPs) as described previously.³⁵² This cohort was selected from the in-house collection of

the Department of Human Genetics of Radboud University Medical Center (Nijmegen, The Netherlands) containing patients with unexplained ID. Candidate loss-of-function mutations as well as highly conserved missense mutations (PhyloP>5) were validated by standard Sanger sequencing approaches on DNA extracted from peripheral blood. For assessing the *de novo* occurrence of validated mutations, DNA from the parents was tested. This study was approved by the institutional review board 'Commissie Mensgebonden Onderzoek Regio Arnhem-Nijmegen'.

Deposition of genotypes and phenotypes in a locus-specific database.

All mutations and phenotypes reported in this publication are deposited in the locus-specific database for WAC, under the realm of the Leiden Open Variation Database (LOVD; <http://databases.lovd.nl/shared/genes/WAC>). Variant information and phenotypes can be retrieved using the following submission entries: Individual 1: #00054835; Individual 2: #00054836; Individual 3: #00054837; Individual 4: #00054838; Individual 5: #00054839; Individual 6: #00054848; Individual 7: #00054849; Individual 8: #00054850; Individual 9: #00054851; and Individual 10: #00054852.

Drosophila lines and maintenance

Fly stocks were kept on standard *Drosophila* diet (cornmeal/sugar/yeast) at 25 °C and 45–60% humidity at 12 h:12 h light/dark cycle. Inducible RNAi lines targeting the *Drosophila* WAC orthologue *CG8949* (*vdrc48307*, *vdrc107328*) and corresponding genetic background control lines (*vdrc60000*, *vdrc60100*) were obtained from the Vienna *Drosophila* RNAi Center.³⁷⁰ The *s19* value, which refers to the specificity of the dsRNA hairpin construct,³⁷⁷ is 1.00 for *vdrc48307* with no off-target and 0.99 for *vdrc107328* (two possible off-targets, *CG11122* and *CG11354*).

RNAi was induced using the UAS-Gal4 system and the panneuronal driver lines: *elav-Gal4 w¹¹¹⁸*; *2xGMR-wIR*; *elav-Gal4*, *UAS-Dicer-2* or *nSyb-Gal4 w¹¹¹⁸*, *UAS-Dicer-2*; *nSyb-Gal4*.^{354; 368} Flies were reared and tested at 25°C (*elav-Gal4*) and 28°C (*nSyb-Gal4*) and 70% humidity. The ubiquitous actin-Gal4 driver *w¹¹¹⁸*; *P(w[+mC]=Act5c-Gal4)/CyO* obtained from Bloomington *Drosophila* Stock Center,³⁷⁸ was used to generate RNAi-mediated knockdown for quantitative PCR (qPCR).

Analysis of CG8949 mRNA levels from larval brains by qPCR

RNA isolations (three biological replicates) from third instar larvae brains were performed using Arcturus Picopure RNA Isolation Kit (Life Technologies, Bleiswijk, The Netherlands). RNA was treated with DNase (DNAfree Kit, Ambion, Bleiswijk, The Netherlands). First-strand cDNA synthesis was performed using iScript cDNA Synthesis Kit (Biorad, Veenendaal, The Netherlands). Gene expression was analysed by real-time PCR (7900HT Fast Real-Time PCR system, Applied Biosystems, Bleiswijk, The Netherlands). PCR reactions were

performed in a volume of 25 µl containing 150 nM primers and GoTag Green Mastermix (Promega, Leiden, The Netherlands). Primer sequences used for amplification of *CG8949*: 5'-TGGAATTACGACAACGATGG-3' and 5'-TAACTGGCTTCCGAGGTAGG-3'. *BetaCop* was used as reference gene, primer sequences: 5'-AACTACAACACCCTGGAGAAGG-3', 5'-ACATCTTCTCCCAATTCCAAAG-3'.

Light-off jump reflex habituation

The light-off jump reflex habituation assay was performed as previously described.³⁶⁹ Habituation of the startle jump response towards repeated lightoff stimuli of 3–7-day-old individual male flies was tested in two independent 16-unit light-off jump habituation systems. A total of 32 flies (16 flies/system) were simultaneously exposed to series of 100 short (15 ms) light-off pulses with a 1 s inter-pulse interval. The noise amplitude of wing vibrations following every jump response was recorded for 500 ms after the start of light-off pulse and an appropriate threshold was applied to filter out the background noise. Data were collected and analysed by a custom-made Labview Software (National Instruments). High initial jumping responses to light-off pulse decreased with the growing number of trials and flies were considered habituated when they failed to jump in five consecutive trials (non-jump criterion). Habituation was scored as the mean number of trials required to reach the non-jump criterion (trials to criterion, TTC). Main effects of genotype (mutant vs control), day and test system on log-transformed TTC values were tested using linear model regression analysis (lm) in R statistical software (R version 3.0.0 (2013-04-03)).³⁷⁹



11

PURA syndrome: clinical delineation and genotype-phenotype study in 32 patients with review of published literature

This chapter has been published as:

Margot R.F. Reijnders, Robert Janowski, Mohsan Alvi, Jay E Self, Ton J. van Essen, Maaïke Vreeburg, Rob P.W. Rouhl, Servi J.C. Stevens, Alexander P.A. Stegmann, Jolanda Schieving, Rolph Pfundt, Katinke van Dijk, Eric Smeets, Connie T.R.M. Stumpel, Levinus A. Bok, Jan Maarten Cobben, Marc Engelen, Sahar Mansour, Margo Whiteford, Kate E Chandler, Sofia Douzgou, Nicola S Cooper, Ene-Choo Tan, Roger Foo, Angeline H.M. Lai, Julia Rankin, Andrew Green, Tuula Lönnqvist, Pirjo Isohanni, Shelley Williams, Ilene Ruhoy, Karen S. Carvalho, James J. Dowling, Dorit L. Lev, Katalin Sterbova, Petra Lassuthova, Jana Neupauerová, Jeff L. Waugh, Sotirios Keros, Jill Clayton-Smith, Sarah F. Smithson, Han G. Brunner, Ceciel van Hoeckel, Mel Anderson, Virginia E. Clowes, Victoria Mok Siu, The DDD Study, Paulo Selber, Richard J. Leventer, Christoffer Nellaker, Dierk Niessing, David Hunt*, Diana Baralle*

Journal of Medical Genetics (2017) 55, 104-113

* These authors contributed equally

Abstract

Background

De novo mutations in *PURA* have recently been described to cause PURA syndrome, a neurodevelopmental disorder characterised by severe intellectual disability (ID), epilepsy, feeding difficulties and neonatal hypotonia.

Objectives

To delineate the clinical spectrum of PURA syndrome and study genotype-phenotype correlations.

Methods

Diagnostic or research-based exome or Sanger sequencing was performed in individuals with ID. We systematically collected clinical and mutation data on newly ascertained PURA syndrome individuals, evaluated data of previously reported individuals and performed a computational analysis of photographs. We classified mutations based on predicted effect using 3D *in silico* models of crystal structures of *Drosophila*-derived Pur-alpha homologues. Finally, we explored genotype-phenotype correlations by analysis of both recurrent mutations as well as mutation classes.

Results

We report mutations in *PURA* (purine-rich element binding protein A) in 32 individuals, the largest cohort described so far. Evaluation of clinical data, including 22 previously published cases, revealed that all have moderate to severe ID and neonatal-onset symptoms, including hypotonia (96%), respiratory problems (57%), feeding difficulties (77%), exaggerated startle response (44%), hypersomnolence (66%) and hypothermia (35%). Epilepsy (54%) and gastrointestinal (69%), ophthalmological (51%) and endocrine problems (42%) were observed frequently. Computational analysis of facial photographs showed subtle facial dysmorphism. No strong genotype-phenotype correlation was identified by subgrouping mutations into functional classes.

Conclusion

We delineate the clinical spectrum of PURA syndrome with the identification of 32 additional individuals. The identification of one individual through targeted Sanger sequencing points towards the clinical recognisability of the syndrome. Genotype-phenotype analysis showed no significant correlation between mutation classes and disease severity.

Introduction

Deletions in the 5q31.2q31.3 region have been reported to cause a syndrome comprising severe developmental delay, profound hypotonia, feeding difficulties, abnormal breathing pattern and seizures.³⁸¹⁻³⁸²⁻³⁸⁵ Within this region, haploinsufficiency of the purine-rich element binding protein A (PURA) has been suggested as responsible for the observed neurodevelopmental features by narrowing the region of overlap to three genes.³⁸⁴ Further evidence for the essential role of PURA in development came from two articles describing 15 individuals with intellectual disability (ID) and *de novo* mutations in *PURA*.^{386; 387} The pathogenicity of mutations in *PURA* was further supported by publication of a cohort of six additional individuals with ID and *de novo* mutations in *PURA* and two case reports.³⁸⁸⁻³⁹⁰

The functionality of PURA is dependent on three PUR motifs, PUR I, PUR II and PUR III.^{391; 392} The ubiquitously expressed PURA protein, Pur-alpha, has multiple regulatory functions in processes such as DNA replication and transcription, mRNA transport and DNA repair.^{392; 393} Several studies have shown that Pur-alpha has an important role in neuronal development and differentiation.³⁹⁴ In animal models, Pur-alpha is involved in postnatal brain development.^{395; 396} The Pur-alpha deficient knockout mouse model has a severe neurological phenotype, including tremor and spontaneous seizures.³⁹⁶ Recently, heterozygous *PURA* knockout mice have been reported to develop gait abnormalities, hypotonia and memory deficits. Immunohistochemical assays displayed a reduced number of neurons in cerebellum and hippocampus of these animals, which is in line with their neurological, behavioural and cognitive phenotypes.³⁹⁷

Mental retardation, autosomal dominant type 31 [MIM616158], or 'PURA syndrome', the name of the syndrome adopted by the parents and clinical community alike, is phenotypically hallmarked by a wide spectrum of neurodevelopmental problems, including severe neurodevelopmental delay, epilepsy, abnormal movements, hypotonia and brain abnormalities. Additionally, neonatal respiratory insufficiency and feeding difficulties have been reported. Outside of this core phenotype, a broad variability in clinical severity has been observed within PURA syndrome.³⁸⁶⁻³⁸⁹ Different types of mutations have been identified so far and mutations occur throughout the gene. It has been suggested that disruption of PUR repeat III possibly results in a more severe phenotype, but additional studies are needed to better understand the genotype-phenotype correlation of *PURA* mutations.³⁸⁷

We present 32 individuals with *de novo* mutations in *PURA*. With this large cohort, we delineate the clinical spectrum of PURA syndrome. Computational modelling shows that subtle overlapping facial dysmorphism is present. Additionally, we present an approach for genotype-phenotype analysis of identified mutations and associated phenotypes, revealing that observed phenotypic variability is probably not associated with type and localisation of mutations.

Results

Identification of 32 individuals

We report 32 individuals with mutations in *PURA* (Figure 11.1). All, but two, were identified using WES in a diagnostic setting (individuals 1–10, 20–27, 30 and 32) or research setting (individuals 12–19, 29 and 31). Individual 28 was identified using targeted MPS of a severe childhood epilepsy gene panel. Individual 11 was clinically suspected to have

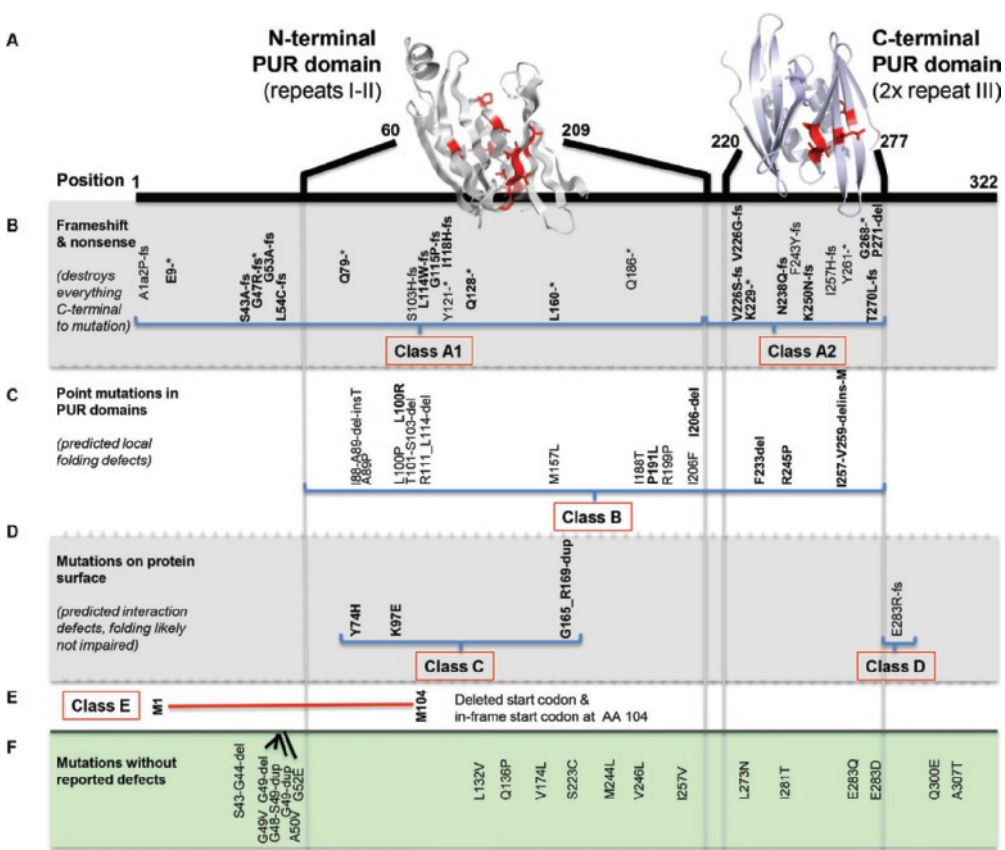


Figure 11.1: Localisation of *PURA* mutations and subdivision in classes. Mutations of individuals identified in our cohort are marked in bold. (A) Homology models of N-terminal and C-terminal PUR domains (grey and blue, respectively) from human Pur-alpha. Residues with single amino acid exchanges are depicted in red with side chains. (B) Location of reported *PURA* frameshift and non-sense mutations. Class A1 mutations are located in the N-terminal PUR domain that affect both N-terminal and C-terminal domains and class A2 mutations occur in the C-terminal domain, affecting only this domain and the C-terminus. (C) Identified point mutations in one of the PUR domains, predicted to cause local folding defects. (D) Four mutations of amino acids located on the protein surface. Three mutations are predicted to affect nucleic acid binding (class C) and one mutation (class D) likely affects a surface-exposed residue, which is not predicted to impair protein folding or nucleic acid binding. This mutation is possibly involved in not yet understood functions such as protein–protein interactions. (E) Deletion (red) caused by mutation p.(Met1?) (class E), which disrupts the start codon. The protein is likely expressed from the next in-frame start codon at amino acid 104, causing loss of a functional N-terminal PUR domain, but has an intact C-terminal PUR domain. (F) Localisation of mutations reported in healthy controls in the ExAC database. All these mutations are predicted to have no effects on the folding and the function of the Pur-alpha protein. PUR, purine-rich element.

PURA syndrome based on post-term birth, severe neonatal hypotonia, hypersomnolence, exaggerated startle, persistent apnoeic episodes requiring continuous SpO₂ monitoring and supplementary oxygen, and feeding difficulties warranting nasogastric tube placement. Targeted Sanger sequencing revealed a *PURA* mutation. Mutations were *de novo* in 29 individuals. For the remaining three individuals, no parental blood was available for testing. None of the individuals had a second, likely pathogenic, genetic variant in WES data. All identified missense mutations localised in one of the PUR repeat sequences (Figure 11.1).

Clinical delineation of 32 individuals

Clinical features observed in individuals reported in this study are summarised in Table 11.1. Numbers and percentages of features were corrected for the number of individuals without available information on a specific feature. More extensive and detailed clinical information is available in Supplemental Table 11.1.

Gestation and neonatal problems

More than half (56%) of individuals were born after >41 weeks' gestation. Induction of labour and/or caesarian section were reported in 14 individuals. In 55% (6/11), excessive hiccups were mentioned in utero. Neonatal problems were evident in all children immediately after birth. Hypotonia was present in all but one of the neonates, often causing feeding difficulties (81%). Apnoeas and congenital hypoventilation were reported in 48% of the neonates. Both feeding and respiratory problems often required monitoring in hospital: in 20 neonates, tube feeding, oxygen supplementation and/or mechanical ventilation were needed. Hypersomnolence (66%) and hypothermia (37%) were observed in a significant number of neonates. In about half of the neonates (58%) an exaggerated startle response was noted.

Development

All individuals with *PURA* mutations have moderate to severe ID with severe language and motor delay. Most individuals remain non-verbal (29/32, 91%), but many have better receptive language than expressive language and can follow simple instructions. Several never achieved independent ambulation (10/24 with age >5 years, 42%). For those who did, the age of first steps ranged from 28 months to 7 years. Regression of achieved skills due to the onset of seizures has been reported in six individuals.

Neurological abnormalities

A variety of neurological problems were observed in individuals with *PURA* mutations. Hypotonia, often more prominent in the trunk, was present from birth. Spasticity of extremities at older age was reported in three individuals, and Babinski response was reported in another two. Gait, if achieved, was often unstable and broad based. Stereotypic hand movements, in some cases such as described for Rett syndrome [MIM312750], were observed in several individuals (36%). Six out of 30 (20%) were mentioned to have a

Table 11.1 Percentage of clinical features reported in individuals in this article (n=32) and meta-analysis with previously published PURA individuals (n=22) ³⁸⁶⁻³⁸⁹

Clinical feature	This article (n = max. 32)		Literature (n = max. 22) ³⁸⁶⁻³⁸⁹		Total (n = max. 54)	
	Percentage	Number	Percentage	Number	Percentage	Number
Growth						
Short stature (≤ -2.5 SD)	19%	5/27	14%	3/22	16%	8/49
Pregnancy/delivery						
Gestational age > 41 weeks	56%	18/32	60%	3/5	57%	21/37
Neonatal problems						
Hypotonia	97%	31/32	94%	15/16	96%	46/48
Feeding difficulties	81%	25/31	73%	16/22	77%	41/53
GORD	28%	8/29	17%	1/6	26%	9/35
Breathing problems	48%	15/31	68%	15/22	57%	30/53
Hypersomnolence	66%	19/29	NR	NR	66%	19/29
Hypothermia	37%	10/27	25%	1/4	35%	11/31
Excessive hiccups in utero	55%	6/11	NR	NR	55%	6/11
Neurological abnormality						
Moderate to severe ID	100%	32/32	100%	22/22	100%	54/54
Hypotonia	97%	31/32	80%	4/5	95%	35/37
Stereotypic hand movements	36%	8/22	NR	NR	36%	8/22
Exaggerated startle response	58%	11/19	27%	4/15	44%	15/34
Epilepsy	50%	16/32	59%	13/22	54%	29/54
Delayed myelination	28%	9/32	24%	5/21	26%	14/53
Movement disorder	20%	6/30	30%	3/10	22%	9/40
Other brain abnormalities	31%	10/32	29%	6/21	30%	16/53
Skeletal abnormality						
Scoliosis	28%	9/32	18%	2/11	26%	11/43
Hip dysplasia	23%	7/31	0%	0/11	17%	7/42
Hyperlaxity	43%	9/21	9%	1/11	31%	10/32
Gastrointestinal abnormality						
Constipation	62%	18/29	50%	3/6	60%	21/35
Drooling	66%	20/30	83%	5/6	69%	25/36
Respiratory abnormality	26%	8/31	27%	4/15	26%	12/46
Cardiac abnormality	13%	4/32	25%	2/8	15%	6/40
Urogenital abnormality	26%	8/31	9%	1/11	21%	9/42
Ophthalmological abnormality	40%	12/30	67%	14/21	51%	26/51
Endocrine abnormality						
Vitamin D deficiency	47%	7/15	25%	1/4	42%	8/19
Other	26%	5/19	50%	2/4	30%	7/23
Skin abnormality						
Soft skin	46%	6/13	NR	NR	46%	6/13

Abbreviations: GORD = gastro-oesophageal reflux disease; ID = intellectual disability; NR = not reported;
PURA = purine-rich element binding protein A

movement disorder, including dystonia, chorea-like movements, seizure-like movements with normal EEG and ataxic movements. Delayed myelination is the most frequently reported brain abnormality observed on MRI images (28%). Other, mostly a-specific, observed brain abnormalities (in 31% of individuals) include white matter abnormalities, prominent periventricular spaces, mild parenchymal atrophy, widening of lateral ventricles and underdeveloped rostrum of the corpus callosum. Peripheral neuropathy was measured in five individuals at young age (Supplemental Table 11.1).

Half of the individuals (50%) were diagnosed with epilepsy. For most of them, seizures started at the age of 2–4 years, but the age of onset ranged from 6 months to 15 years. Different seizure types were reported, including generalised tonic-clonic seizures, focal seizures, absence seizures, epileptic spasms, tonic seizures, drop attacks and over time, evolution to Lennox-Gastaut syndrome. The epilepsy is often refractory to medical treatment and some individuals have never been seizure free. A longer follow-up period will be required to determine the true prevalence of epilepsy, how refractory it is and what treatments are most effective.

Ophthalmological abnormalities

Eye abnormalities and visual problems were described frequently in PURA syndrome individuals (40%). Strabismus (often oesotropia) was the most commonly reported finding along with strabismus-associated refractive errors (often hypermetropia). Cortical visual impairment was reported in three individuals; however, it is important to note that diagnostic criteria for this condition are highly variable. One individual was treated successfully for a congenital retinoblastoma. Nystagmus was also described in a small number of cases (4/29; 14%) but details of the phenotype are missing in reported cases.

Skeletal abnormalities

Progressive hip dysplasia and scoliosis were present in respectively 23% and 28% of the cases and could possibly be related to chronic truncal hypotonia (with or without abnormalities of limb tone), joint laxity and delayed or incomplete motor development. Surgical correction was required in five individuals in this cohort to date. The oldest individual was 18 years at time of surgery. Longer follow-up might show a larger number of surgical interventions, since the majority of included individuals (27/32) had an age of examination below 18 years. Individuals showed a tendency to short stature with five having a height ≤ 2.5 SD, and six others with a height between -2 and -2.5 SD. From the remaining 16 individuals, only 2 had a height >0 SD. Low bone density (z-score -4 ; 57%) was identified in a single case. Pes planus was present in 11 individuals.

Gastrointestinal abnormalities

Due to severe hypotonia, swallowing problems and excessive drooling were observed in

more than half of the individuals (66%). Constipation was often present (62%) and may be severe, requiring the use of laxatives from early age.

Endocrine abnormalities

A diverse range of endocrine problems were reported, with low vitamin D most commonly reported (47%), and less frequently observed abnormalities including aberrant sex hormone levels, abnormal cortisol response and aberrant thyroid hormone levels (26%) (Supplemental Table 11.1). It is likely that these problems are more common than observed, since not all individuals have been tested for endocrine abnormalities.

Congenital structural malformations

Although not frequently reported, some individuals had congenital malformations of the heart, urogenital and/or respiratory tract. Cardiac malformations included a ventricular septal defect in two individuals, and an aberrant left subclavian artery, pulmonary stenosis and mild, spontaneously closed persistent ductus arteriosus each in one individual. Abnormalities of the urogenital tract were described in five individuals: cryptorchidism in three, kidney stones in two, and congenital hydronephrosis with megaureter and urinary reflux each in one individual. Laryngeal cleft was reported in a single case.



Figure 11.2: Photographs of 21 individuals with PURA mutations.

Shared facial dysmorphism includes high anterior hairline, almond-shaped palpebral fissures, full cheeks and hypotonic face. Strabismus is present in several of the PURA (purine-rich element binding protein A) individuals. Additionally, independent dysmorphologists also observed eversion of lower lateral eyelids, prominent, well-defined philtrum and retrognathia in a subset of the individuals.

Facial dysmorphisms

Photographs of 21 individuals are shown in Figure 11.2. No striking similarities were observed, but a myopathic face, high anterior hairline, almond-shaped palpebral fissures and full cheeks were reported recurrently by clinicians. The presence of these features was confirmed by a panel of five independent dysmorphologists, all being asked to analyse photographs of PURA syndrome individuals from Figure 11.2. They also mentioned eversion of lower lateral eyelids, prominent, well-defined philtrum and retrognathia as features present in a subset of individuals.

Expansion of the cohort with 22 previously published cases

We structurally analysed mutational and clinical data of 22 previously published individuals.³⁸⁶⁻³⁸⁹ An overview of published mutations is present in Figure 11.1 and summarised clinical information is available in online Supplemental Table 11.2.

Recurrence of mutations

We found four recurrent mutations: p.(Lys97Glu) in individual 20 and subject #4 in Lalani *et al*³⁸⁶; p.(Phe271del) in individual 28 and subject #1 in Lalani *et al*³⁸⁶; p(Arg245Pro) in individuals 1 and 15; p.(Phe233del) in individuals 4, 5 and 14, patient 4 in Tanaka *et al*³⁸⁸ and patient 4 in Hunt *et al*.³⁸⁷ In addition, individuals 24 and 31 had two distinct mutations (c.153delA and c.155delG, respectively) that were both predicted to give rise to the same truncated protein product (p.(Leu54Cysfs*24)). There were six individuals between whom two distinct missense mutations were identified at each of three different amino acid positions: p.(Leu100Arg) and p.(Leu100Pro) in individual 21 and subject #1 in Lalani *et al*;³⁸⁶ p.(Ile206del) and p.(Ile206Phe) in individual 25 and patient 3 in Hunt *et al*;³⁸⁷ p.(Val226Serfs*68) and p.(Val226Glyfs*67) in individuals 6 and 10.

Clinical evaluation

Clinical features such as neurodevelopmental delay, epilepsy, and brain and ophthalmological abnormalities have been reported in all previously published cohorts. But for the majority of features, no clinical information is available across all previously published cohorts. For example, hypersomnolence, stereotypic hand movements and excessive hiccups in utero have not been reported in literature before. Using the summarised clinical information (Supplemental Table 11.2), we calculated the frequency of clinical features in all 54 individuals (Table 11.1).

Computational analysis of facial dysmorphism

Since no striking overlap in facial dysmorphism was observed, but similarities between several individuals have been reported by clinicians, we performed an objective computational analysis in this paper presented and previously published photographs, and compared it to the average image based on healthy controls (Figure 11.3). Ages

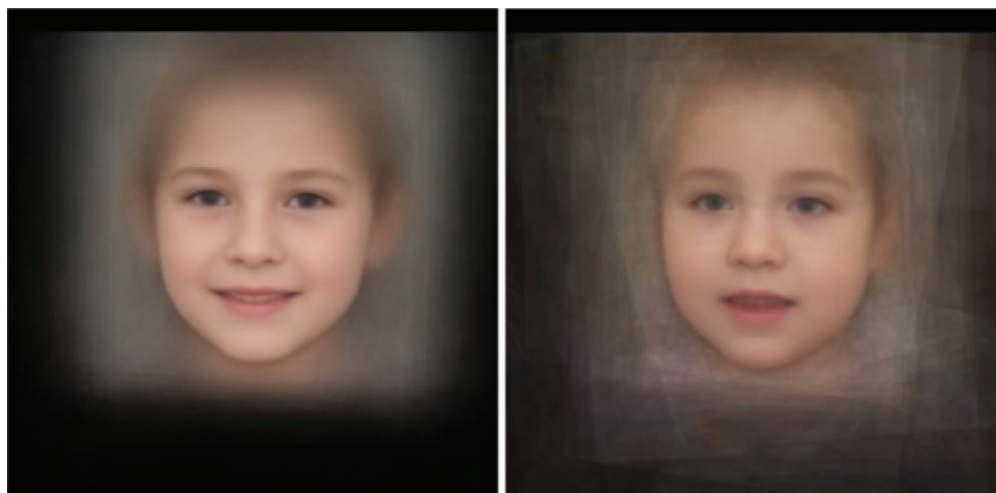


Figure 11.3: Computational analysis of photographs of PURA syndrome individuals.

Result of an objective computational analysis on photographs of 34 individuals with PURA mutations (ages at photographs ranging from 2 months to 19 years; right image) compared with the average image based on 301 age-matched, healthy controls (left image). The computational modelled PURA face showed a hypotonic face with typically open mouth appearance and full cheeks. Additionally, two independent dysmorphologists reported a slightly abnormal shape of the eyes as (1) shorter palpebral fissures and (2) eversion of lower lateral eyelids. The high anterior hairline observed in a subset of individuals is not visible on this computational model of the PURA face.

at photographs ranged from 2 months to 19 years. A panel of five independent dysmorphologists agreed that a myopathic face with typically open mouth appearance and full cheeks was visible on the computational modelled PURA face. Additionally, two dysmorphologists reported a slightly abnormal shape of the eyes as (1) shorter palpebral fissures and (2) eversion of lower lateral eyelids. The high anterior hairline observed in a subset of individuals (6,7,8,9,11,13,14,16,17,28) was not visible on the computational model of the PURA face.

Modelling of all reported mutations into structural homology models of human Pur-alpha

Using the previously reported crystal structures of the N-terminal and C-terminal PUR domains from the fruit fly *D. melanogaster*^{391; 392} as templates, we generated structural homology models for the corresponding domains of the human Pur-alpha paralogue. Based on these homology-modelled structures, all 46 reported mutations were analysed and classified according to their predicted effects on domain folding and function.

The identified frameshift and non-sense mutations affecting the entire protein C-terminal to the site of mutation were defined as class A mutations. Depending on whether mutations occurred in the N-terminal PUR domain and thus affect N-terminal and C-terminal PUR domains or in the C-terminal PUR only affecting this very domain plus C-terminus, we

subclassified them as A1 and A2 mutations, respectively (Figure 11.1B). Mutations that were predicted to affect folding only locally, for instance, via amino acid exchange of a buried residue or short amino acid deletions in a folded region, were defined as class B mutations (Figure 11.1C). Since the RNA and DNA-binding mode of *Drosophila* Pur-alpha has been structurally analysed,³⁹² the human homology model could also be used to predict mutations that affect nucleic acid binding. Such mutations were found in four individuals and were defined as class C mutations (Figure 11.1D). We also found one mutation likely affecting a surface-exposed residue that was neither predicted to impair protein folding nor nucleic acid binding. We suggest that this residue is involved in not yet understood functions such as protein-protein interactions and we defined this as class D mutation (Figure 11.1D). Finally, we classified one mutation as class E: p.(Met1?), causing loss of the start codon. Translation is therefore likely to start at the next in-frame start codon, which is located at protein position 104. As a result, the protein will lack part of the N-terminal PUR domain, while it will have an intact C-terminal PUR domain (Figure 11.1E). As control, we analysed variants reported by ExAC, a database containing variants of healthy controls (Figure 11.1F). All of these variants resulted in single amino acid exchanges that are either located outside the globular domains or were exchanged against residues with similar side chain properties. Variant p.(Gln136Pro) is the only exception to this rule. It is located in the large loop region between repeats I and II and thus likely allows to compensate the changes in backbone conformation upon mutation into proline.

Genotype-phenotype analysis

Remarkable clinical variability has been observed within the total cohort of PURA syndrome individuals. To assess whether this variability could be explained by the type and/or localisation of the identified mutations, we compared (1) phenotypes of individuals with the same mutation and (2) phenotypes of individuals with a mutation of the same class.

Recurrent mutations

We observed four recurrent mutations and two mutations with a similar protein effect within the cohort, and compared their phenotypes. For all mutations, remarkable differences between individuals with the same mutation were present (Table 11.2). For example, drug-resistant epilepsy was present in three out of five individuals with mutation p.(Phe233del) and a cardiac abnormality in two out of five individuals. The older of the two individuals without epilepsy could walk, in contrast to the three individuals with epilepsy, who were not able to walk or speak. For mutation p.(Leu54Cysfs*24), individuals had more features in common, such as epilepsy, scoliosis and short stature, but variation in ability to walk and heart abnormality was also observed.

Comparison of mutation classes

We compared phenotypes of individuals subdivided by mutation class (Supplemental

Table 11.2 Recurrent *PURA* mutations and phenotype of affected individuals

Mutation	Individual/Age	Reference	Phenotype
p.(Leu54Cysfs*24)	Individual 24 17 years	This article	Hypotonia, feeding difficulties requiring TF, GORD, apnoeas, hypersomnolence, hypothermia, severe ID, no speech, walking at age 4 years, short stature (-5 SD), epilepsy, mild PDA, mild strabismus, scoliosis, hip dysplasia, small hands and feet.
	Individual 31 16 years	This article	Hypotonia, feeding difficulties requiring TF, hypersomnolence, severe ID, no speech, not able to walk, short stature (-2.75 SD), epilepsy, scoliosis, epilepsy, long fingers
p.(Lys97Glu)	Individual 20 2.5 years	This article	Hypotonia, mild feeding difficulties, hypersomnolence, no speech, not able to walk, mild constipation, pes planus
	Subject #4 21 months	Lalani et al ³⁸⁶	Hypotonia, feeding difficulties, respiratory difficulties, seizures, ID, non-verbal, non-ambulatory, short stature (1%), strabismus, duplex left kidney, hydronephrosis
p.(Phe233del)	Individual 4 14 years	This article	Hypotonia, feeding difficulties, apnoeas, hypersomnolence, severe ID, no speech, able to walk, ASD, stereotypic hand movements, delayed myelination, swallow problems, severe hypermetropia, strabismus, low vitamin D, low ferritin, pes planus, abnormal peripheral nerve testing
	Individual 5 19 years	This article	Hypotonia, feeding difficulties, apnoeas, hypersomnolence, hypothermia, severe ID, no speech, not able to walk, stereotypic hand movements, epilepsy, nystagmus, delayed myelination, drooling, constipation, strabismus, CVI, scoliosis, hip dysplasia, low bone mineralization, low vitamin D, anaemia, delayed puberty, small hands, pes planus
	Individual 14 9 years	This article	Hypotonia, severe ID, no speech, first steps age 7 years, epilepsy, delayed myelination, aberrant left subclavian artery, VSD, drooling, refraction abnormality, strabismus, low vitamin D, high cholesterol
	Patient 4 6 months	Tanaka et al ³⁸⁸	Hypotonia, CVI, periventricular leukomalacia
	Patient 4 6 years, 9 months	Hunt et al ³⁸⁷	Hypotonia, feeding difficulties requiring TF, apnoeas, hypothermia, severe ID, not able to walk, essentially non-verbal, exaggerated startle response, dystonia, dyskinesia, epilepsy, CVI, delayed myelination, excessive extra-axial fluid spaces, cerebral atrophy, hyperprolactinaemia, blunted cortisol stress response, low vitamin D
p.(Arg245Pro)	Individual 1 19 years	This article	Hypotonia, feeding difficulties, severe ID, no speech, not able to walk, autistic-like traits, chorea-like movements, Babinski response, delayed myelination, scoliosis, low vitamin D, hypogonadotropic hypogonadism
	Individual 15 9 years	This article	Hypotonia, feeding difficulties requiring TF, apnoeas, hypersomnolence, hypothermia, severe ID, no speech, first steps at age 4 years, epilepsy, drooling, constipation, pes planus
p.(Phe271del)	Individual 28 11 months	This article	Hypotonia, neonatal convulsions, moderate ID
	Subject #1 6 months	Lalani et al ³⁸⁶	Hypotonia, feeding difficulties, respiratory difficulties, seizures, ID, pedal edema

Abbreviations: ASD = autism spectrum disorder; CVI = cortical visual impairment; GORD = gastro-oesophageal reflux disease; ID = intellectual disability; PDA = persistent ductus arteriosus; TF = tube feeding; VSD = ventricular septum defects

Table 11.3). Overall, no significant differences in percentage of observed features between mutation classes were present. One exception is mutation p.(Glu283Argfs*45), which we classified as class D. This mutation affects only the very C-terminus (all three PUR domains are intact) and it is unlikely that this mutation interrupts protein folding or nucleic acid binding (Supplemental Table 11.4; Figure 11.1). Clinical features in this individual are less severe than observed in other individuals: the 14-year-old girl could walk independently from the age of 2 years and could speak in sentences. Furthermore, no neonatal problems, hypotonia, epilepsy or brain abnormalities were present.³⁸⁷ The clinical presentation of this girl is remarkably milder than other individuals, and could possibly be explained by the localisation of the mutation at the C-terminus.

Discussion

We report 32 individuals with PURA syndrome, the largest cohort so far. We systematically collected mutation and clinical data on these and previously published individuals. With information on a total of 54 individuals, we were able to delineate the clinical spectrum of this recently discovered syndrome.

The detailed and systematic collection of phenotypic data allowed us to observe less frequently reported features related to PURA syndrome. Clinical problems such as excessive hiccups in utero (55% in our cohort), hypersomnolence (66% in our cohort), stereotypic hand movements (36% in our cohort), regression since onset of seizures (38% of individuals with epilepsy in our cohort), soft skin (46% in our cohort) and short stature (19% in our cohort) have not been reported before in published studies. Post-term birth (56% in our cohort), gastro-oesophageal reflux disease (28% in our cohort), hypothermia (37% in our cohort), hypotonia after the neonatal period (97% in our cohort), constipation (62% in our cohort) and endocrine abnormalities (47% in our cohort) only have been mentioned in one other study. Similar to many other recently discovered novel syndromes for which only a small subset of individuals have been described initially, frequencies of features related to PURA syndrome were difficult to estimate based on reported numbers. By expanding the cohort, we were able to report a significant number of features not formally associated with PURA syndrome before. These features could be helpful for diagnosis and management of PURA syndrome individuals.

The phenotype of PURA syndrome has been reported as difficult to recognise in daily clinical practice. Here, we show with the identification of one individual using targeted Sanger sequencing in a neonate, that it is possible to recognise PURA syndrome based on clinical features. Severe neonatal problems including hypotonia, respiratory and feeding difficulties, hypersomnolence and hypothermia, in the majority combined with post-term birth, could point towards the early diagnosis of PURA syndrome. Although not a consistent feature, the presence of myoclonic jerks in a neonate with other features of PURA syndrome may

be a useful feature to distinguish from Prader-Willi syndrome or peripheral neuromuscular disorders such as spinal muscular atrophy (SMA). Known syndromes such as congenital hypoventilation syndrome [MIM209880], Prader-Willi syndrome [MIM176270], Rett syndrome, Angelman syndrome [MIM160900], myotonic dystrophy [MIM160900] and SMA [MIM253300] have been tested often in PURA syndrome individuals. We recommend to consider PURA syndrome in the differential diagnosis of any child tested for one or more of the above syndromes, especially in the context of the aforementioned neonatal problems and/or unexplained developmental delay.

After diagnosis of PURA syndrome, careful management of existing medical problems and evaluation of health issues associated with PURA syndrome is essential. Our phenotypic overview shows that different organ systems could be affected. Therefore, we recommend routine care in a multidisciplinary team, with at least a (child) neurologist, ophthalmologist, orthopaedic surgeon and paediatrician for children aged below 18 years involved. Epilepsy, present in half of the cohort, is a major problem. Some individuals have never been seizure free and regression of achieved developmental milestones since onset of seizures has been reported. In future, studies focusing on epilepsy will be important to optimise treatment of seizures related to PURA syndrome.

In general, individuals with epilepsy were more severely affected than individuals without epilepsy. However, this explained only partially the variation in developmental phenotypes observed between individuals. To assess whether the clinical variability could be related to the type and localisation of the identified mutations, we performed a genotype-phenotype study by analysing recurrent mutations and classifying mutations based on a homology model derived from a crystal structure of the *Drosophila* PUR-alpha homologue. Interestingly, remarkable differences, such as the presence or absence of epilepsy or congenital malformations, were observed between individuals with the same mutation. For the remaining mutations, class D mutations with localisation at the very end of the C-terminus could possibly be correlated with a less severe phenotype. But for the other mutations, phenotypic variability and severity could not be related to the localisation and type of mutation. Based on these results, we suggest that other genetic and biological mechanisms contribute to the phenotypic variability. Further studies are needed to obtain insight into these underlying pathological mechanisms.

With the identification of 54 individuals within 2.5 years of the initial description of the condition, PURA syndrome appears to be a relatively frequent cause of ID. Previously, frequencies between 0.3% and 0.5% in neurodevelopmental cohorts have been reported.³⁸⁶⁻³⁸⁸ Within our cohort, 10 individuals were identified in the diagnostic lab facility of the Radboudumc, Nijmegen. These were found in a total of 4700 individuals (~0.2%) with ID referred for WES. This frequency is probably an underestimate. It has been reported

that *PURA*, a single-exon gene rich in guanine-cytosine content, could be challenging to capture, similar to the first exon of multiexonic genes.¹⁰⁶ Furthermore, all 10 mutations have been identified after the introduction of *PURA* as an ID gene in diagnostic gene panels at the end of 2014, while WES had already been in use in the diagnostic lab of the Radboudumc from 2013. Therefore, to more accurately gauge the frequency within this population, a targeted re-evaluation of all available exome data would be necessary.

In conclusion, we present a detailed overview of the clinical spectrum of *PURA* syndrome and a computational modelled *PURA* face, which is essential for better recognition of this newly recognised disorder and surveillance and management of symptoms after initial diagnosis. Our suggested approach for genotype-phenotype analysis by modelling of identified *PURA* mutations into human homology models of experimentally determined crystal structures from *Drosophila* Pur-alpha, showed that clinical variation observed between individuals is probably not related to type and localisation of mutations, but should be sought in other genetic and biological mechanisms.

Material and methods

Identification of PURA syndrome individuals and collection of clinical and mutational data

Whole exome sequencing (WES) was performed in individuals with ID in either diagnostic ($n=20$) or research ($n=10$) settings (Supplemental Methods). In one individual, targeted massively parallel sequencing (MPS) of a severe childhood epilepsy gene panel was used, and in a second individual, clinically suspected to have *PURA* syndrome, *PURA* [NM_005859] was investigated with targeted Sanger sequencing. Parental samples were tested to assess *de novo* state. A questionnaire with clinical and mutational details was completed by clinicians. Written consent for publication of photographs was obtained from legal guardians.

Evaluation of previously published cases

Twenty-two individuals with *PURA* mutations have been described in four different publications.³⁸⁶⁻³⁸⁹ We used the same questionnaire to collect clinical and mutation data from individuals ($n=22$) reported in the literature.

Computational analysis of facial photographs

To visualise the characteristic facial features of *PURA* syndrome individuals, we generated a realistic, deidentified average based on 34 photographs of individuals (in this paper presented and previously published photographs) at different ages. A control average was created from 301 age-matched, healthy individuals. We used a fully automated algorithm that (1) annotated a constellation of 68 facial landmark points on the face, (2) created an average face mesh and (3) warped faces of each individual onto the average face mesh. The face averaging algorithm was inspired by previous work^{398; 399} and improved upon with better

deidentification of individuals (Supplemental Methods). Interpretation of the composite PURA face was performed by a panel of five independent dysmorphologists.

Functional assessment of PURA mutations in structural homology models

For homology modelling of mutations, the crystal structures of the N-terminal PUR domain (PDB-ID: 3K44) and of the C-terminal PUR domain (PDB-ID: 5FGO) from *Drosophila melanogaster* were used as template. Homology structures of human PUR domains were calculated using the PHYRE2 program.⁴⁰⁰ In the case of the C-terminal PUR domain, the sequences of two repeat ILLs were merged into one peptide chain to allow for modelling of the entire domain. The human homology models showed high confidence scores and were used for in silico mutational analyses using the program Coot.⁴⁰¹ All mutations identified in our cohort and previously published cohorts³⁸⁶⁻³⁸⁹ were included. Variants in *PURA* reported in the Exome Aggregation Consortium (ExAC) database⁷¹ were included as controls. Deletions, frameshift, non-sense and missense mutations were manually classified depending on the position, size and function of the affected amino acids. Point mutations were in silico introduced using the standard rotamer library of Coot and analysed for the appearance of steric clashes or repulsive forces of side chains. They were classified based on the orientation of mutated side chains and on our knowledge of functionally important surface regions. Figures were prepared using the program Pymol (www.pymol.org).



12

General discussion and future perspectives

12.1 Neurodevelopmental disorders: a next generation

12.1.1 Next generation NDDs & genetics: discovery of rare NDDs

In 2010, the first gene for a Mendelian disorder was discovered using next generation sequencing.⁴⁰² NGS techniques have steadily improved, and large numbers of novel genes and NDD syndromes, denoted here as ‘next generation NDDs’, have been discovered.¹⁰¹ Currently, ~30% of all NDD patients referred for WES analysis can be conclusively diagnosed by molecular analysis of ‘known NDD genes’ (**Chapter 1**). Yet, a significant proportion of patients remains undiagnosed. For these patients, it is important that identification of novel NDD genes continues. In this thesis, I aimed to discover rare novel NDDs by (1) WES testing of patients with unexplained NDDs on large scale, (2) using different approaches to interpret WES data, and (3) inclusion of patients with mild ID, a less frequently investigated subset of patients (Figure 12.1).

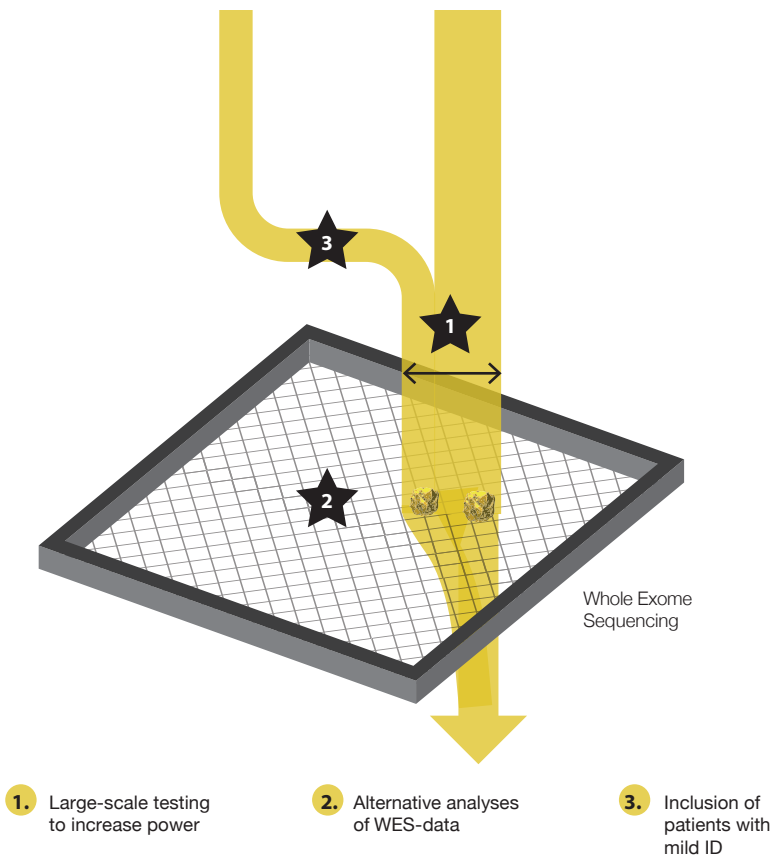


Figure 12.1: Discovery of rare NDDs using three different strategies. To discover rare NDDs, a large cohort of patients (yellow arrow) is tested by WES (the sieve), and mutations causing novel NDDs (gold nuggets) can be filtered out using different approaches to interpret WES data.

By using these strategies, 14 genes were added to the NDD gene list (Table 1): *DLG4*, *PPM1D*, *SMAD6*, *SOX5*, *SYNCRIP*, *TCF20* and *TRIP12* as candidate NDD genes (**Chapter 3**) and *SON*, *TLK2*, *RAC1*, *RHEB*, *USP9X*, *RAB11B* and *WAC* as novel NDD genes. (**Chapter 4-10**) The identification and publication of candidate and novel NDD genes is important and enables the initiation of follow-up studies. In this thesis, I showed that three initial candidate genes could be confirmed as novel genes in additional studies (*RAC1*, *SON* and *TLK2*; **Chapter 4-6**). Three other candidate genes (*PPM1D*, *TCF20* and *TRIP12*) have in the meantime been confirmed as ID/NDD genes by others.⁴⁰³⁻⁴⁰⁵ For both *TRIP12* and *USP9X*, clinical follow-up studies were published recently (Table 12.1).^{406; 407}

Table 12.1 Overview of next generation (candidate) NDD genes described in this thesis (in yellow), with confirmation in literature (in grey).

Gene	Candidate NDD gene		Novel NDD gene		Follow-up NDD gene	
	This thesis	Literature	This thesis	Literature	This thesis	Literature
<i>DLG4</i>	Chapter 3					
<i>PPM1D</i>	Chapter 3			Jansen et al. ⁴⁰³		
<i>RAC1</i>	Chapter 3		Chapter 6			
<i>SMAD6</i>	Chapter 3					
<i>SON</i>	Chapter 3		Chapter 4	Tokita et al. ⁴⁰⁸		
<i>SOX5</i>	Chapter 3	Nesbitt et al. ⁴⁰⁹				
<i>SYNCRIP</i>	Chapter 3					
<i>TCF20</i>	Chapter 3	Babbs et al. ⁴¹⁰		Schafgen et al. ⁴⁰⁴		
<i>TLK2</i>	Chapter 3		Chapter 5			
<i>TRIP12</i>	Chapter 3	O'Roak et al. ⁴¹⁴		Bramswig et al. ⁴⁰⁵		Zhang et al. ⁴⁰⁶
<i>RHEB</i>			Chapter 7			
<i>USP9X</i>		Brett et al. ³⁰⁹	Chapter 9			Au et al. ⁴⁰⁷
<i>RAB11B</i>			Chapter 8			
<i>WAC</i>		Hamdan et al. ³⁵⁵ De Ligt et al. ²⁸	Chapter 10	DeSanto et al. ⁴¹¹		

Increasing power to detect rare NDDs

In the absence of a clinically distinct phenotype, discovery of novel dominant disorders is difficult. Different strategies have been proposed to find more genes involved in the pathogenesis of these non-specific syndromes. One of these strategies is based on the search for different patients with recurrence of mutations within the same gene.⁴¹² However, recurrence of mutations is difficult to achieve if patient cohorts are small and the prevalence of disorders is low. Next generation NDDs are almost always extremely rare. *ARID1B* and *ANKRD11*, causing Coffin-Siris syndrome and KBG syndrome respectively,^{413; 414} are among the most frequently mutated genes in undiagnosed patients referred for WES.³³ Yet, together they explain only for ~1.6% of patients the neurodevelopmental delay. Many other NDDs have an even lower prevalence. For example, mutations in *KMT2D* (Kabuki syndrome)⁴¹⁵

account for ~0.19% of diagnoses, and mutations in *SETBP1* (Schinzel-Giedion)⁴¹⁶ are even rarer with a prevalence of ~0.07%.³³ Inspired by other large studies investigating patients with ASD,^{31; 417} we collected for the studies reported in this thesis a cohort of 826 patients with unexplained ID, a well-investigated subtype of NDDs, in an attempt to identify novel ultra-rare NDDs. In our cohort, we were able to find a mutation in a known NDD gene in 28,6% of patients. Additionally, a significant proportion of patients (386/826, 46.7%) had at least one *de novo* mutation in a gene not previously associated with ID. Several of those genes were mutated in more than one patient in our cohort: the 628 *de novo* mutations comprised only 586 genes. Genes with recurrence included the novel genes *TLK2*, *RAC1*, *RHEB*, and *USP9X*, presented in this thesis, which each occurred at a frequency of 0.24%. Interestingly, even lower frequencies of *de novo* mutations in these genes were observed in the largest NDD cohort published so far consisting of 4,293 patients (0.09% - *TLK2*, 0.07% - *RAC1*, 0% - *RHEB*, and 0.16% - *USP9X*),³³ showing that these novel NDD syndromes are ultra-rare. The chance that recurrence of mutations in such genes occurs, decreases when testing smaller cohorts of patients, and is heavily dependent on the gene-specific mutation rate. Further discovery of these ultra-rare novel disorders thus requires much larger datasets and combined datasets in meta-analysis, as shown in **Chapter 3**.

Alternative analysis of WES data to reveal candidate NDD genes

A large cohort size not only increases the chance to identify recurrence for mutations in the same gene, but also allows alternative analyses on WES data to find novel NDD genes and additional evidence for the pathogenicity of variants. In 2014, Samocha et al. developed a statistical framework, along the line of previous metrics such as the residual variation intolerance score (RVIS),¹²⁴ to investigate excesses of *de novo* loss-of-function and missense mutations per gene and per gene-set. They hypothesized that genes associated with disease would contain more *de novo* mutations than expected by chance, given a certain cohort size.¹¹⁵ In their study, a lack of power was present, since only limited WES data of both control and patient cohorts were available. The Deciphering Developmental Disorders (DDD) study was the first to show that gene-specific mutation rates were helpful for the detection of genes involved in NDDs by using large number of patients ($n \sim 4,000$) and controls.³² In the DDD study, the diagnostic yield was increased with 3% by the discovery of 12 novel NDD genes that showed significantly more *de novo* mutations than expected. In this thesis, we were able to further increase the number of candidate genes associated with ID/NDD by 10, using a similar approach and by adding data of additional patients from our cohort to previously available datasets (**Chapter 3**). The immense power of this methodology is further underscored by the fact that from these 10 candidate genes we described, six have already been confirmed as novel NDD genes in articles including specific phenotypes and/or functional data. But limited cohort size can still be a limiting factor and combined with the rigid statistical correction for multiple testing performed, it might be expected that genes with p-values only slightly below the significance threshold may still be true novel NDD genes.

In our study, two (*CSNK2A1* and *KCNQ5*) out of the first four borderline significant genes, *CSNK2A1*, *CHD3*, *CNOT3*, and *KCNQ5*, have been confirmed as NDD genes by others.^{418;}

⁴¹⁹ The two other two genes, *CHD3* and *CNOT3*, remain interesting candidate genes, given their significant intolerance scores (z-scores=7.15 and 3.89, pLI scores = 1.00 for *CHD3* and *CNOT3*, respectively) and large number of patients with NDD and *de novo* mutations in these genes in the *de novo* database (13 for *CHD3* and 11 for *CNOT3*).²⁵² Clearly, their confirmation as novel NDD gene awaits publication; in fact, such a manuscript was recently submitted for *CHD3* [L. Snijders Blok, et al. personal communication].

One appealing approach, already proposed many years ago,⁴²⁰ is to study not all genes as separate entities, but to evaluate the contribution of functionally related genes in gene-sets (one-entity). So far, a search for mutations and novel NDD genes in biological pathways has been performed mainly on cohorts of patients with specific syndromic phenotypes. Examples include targeted sequencing of ciliary genes in Joubert syndrome patients,⁴²¹ genes in the mTOR pathway in patients with focal cortical dysplasia or megalencephaly,²³³ chromatin-modifiers in patients with features within the Kleefstra phenotypic spectrum,⁴²² and genes interacting with known RASopathy genes in patients suspected to have Noonan syndrome.⁴²³ With the current availability of large WES data sets, in particular WES data linked to HPO terms describing the patient's phenotype,⁴²⁴ time-consuming selection and inclusion procedures to compile specific patient cohorts can be diminished. In **Chapter 7**, I used a pathway approach combined with statistical analyses to discover (possible) pathogenic mutations in our WES data set of 826 patients. Mutations in five candidate NDD genes related to mTOR were present and statistical enrichment analysis showed that significantly more *de novo* mutations were present than expected by chance. Further analysis showed that in aggregate, patients with a mTOR pathway *de novo* mutation had an increased head size, which is in agreement with previous phenotypic observations in patients with mutations in other genes from this pathway.^{190; 232; 233} Based on these results, it is likely that the identified mutations are disease-causing, although individual confirmation of each mutation is still required to unambiguously define them as pathogenic, similar to what we did for mutations in *RHEB*. It would be interesting to perform a similar analysis on different pathways linking specific clinical features that have already been reported to occur with gene mutations in selected pathways. For example, cardiac abnormalities for RASopathy genes, and cerebellar abnormalities for ciliary genes. The feasibility of this approach clearly depends on the availability of detailed phenotype data.

Gene discovery in patients with mild ID

Patients with severe phenotypes are most likely to be referred for genetic testing, since the diagnostic yield is historically considered to be much higher for patients with moderate or severe ID than for patients with mild ID.¹⁰¹ For example, examining subtle structural rearrangements as cause for unexplained ID, revealed a diagnosis in 7.4% of patients with

moderate/severe ID, and only in 0.5% of patients with mild ID.⁴²⁵ For WES, the diagnostic yields of different severity levels are less well documented.¹⁰¹ In the cohort of 826 patients described in this thesis (**Chapter 2**), 219 patients (28%) had mild ID. A *de novo* mutation in a known NDD was present in 61/219 (27.9%) (Figure 12.2). A comparable percentage ($p = 0.924$; Pearson's chi-square test) was present for patients with more severe ID: in 104 patients with moderate or severe ID (27.2%) a *de novo* mutation in a known NDD gene was identified (Figure 12.2). In this thesis, mutations in genes such as *TLK2*, *WAC* and *USP9X* were mainly present in patients with mild ID. These results demonstrate that WES should not be restricted to patients with more severe phenotypes, and that also for patients with mild ID, novel syndromes can be discovered.

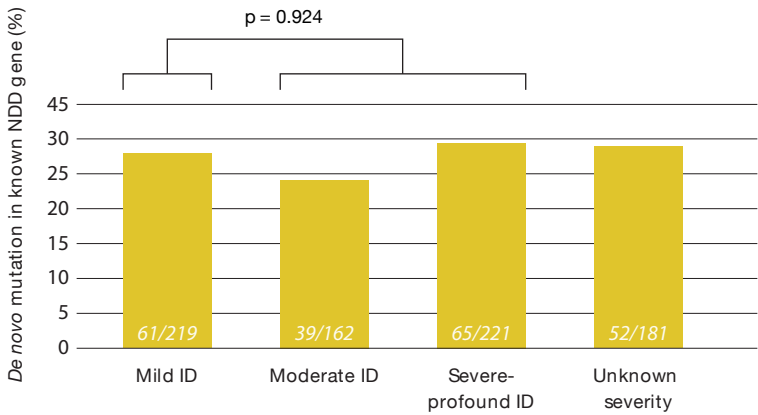


Figure 12.2: Percentage of patients identified with a *de novo* mutation in a known NDD gene in the cohort of 826 patients described in this thesis. Percentages are, in contrast to what has been published for other genetic tests, comparable for different severity levels of ID ($p=0.924$, Pearson's chi-square test).

The majority of studies aiming to identify potentially pathogenic mutations in patients with NDDs using NGS, including the study reported in **Chapter 3**, have focused on the search for *de novo* mutations.^{27-33; 417} This strategy has been developed based on ‘the genetic paradox’: *de novo* mutations can explain theoretically why severe disorders such as NDDs remain relatively frequently in the population, despite negative reproductive selection.^{426; 427} The success of this *de novo* approach is demonstrated by the large number of novel NDDs that have been identified over the last years. By the same token, mild phenotypes in the patients give less reproductive disadvantages, resulting in transmission of a pathogenic mutation to children. In **Chapter 5**, I present an example of a next generation NDD that is not only caused by *de novo* mutations in *TLK2*, but can also develop as a result of an inherited mutation. This has been reported in literature as well for other NDDs such as KBG syndrome (*ANKRD11* mutations), Noonan syndrome, Weaver syndrome (*EZH2* mutations), and Kabuki syndrome (*KMT2D* mutations).⁴²⁸⁻⁴³¹ So, in particular if (one of the) parents have a history of

learning difficulties or behavioral problems and cognitive deficits are mild, there is a higher chance that a Mendelian genetic cause is inherited from a mildly affected parent who was previously undiagnosed. Structured collection of clinical data not only from the patient but also the parents is therefore essential.

12.1.2 Next generation NDDs & clinic: confirmation and delineation by matchmaking

Next generation NDDs have at least one characteristic in common: due to their rarity, and the novelty of the syndromes, limited clinical information is available. Most publications reporting the identification of candidate NDD genes based on large cohort analyses, lack detailed phenotypic description of patients.^{30-32; 417} But the identification of mutations in recurrent genes needs more support than statistics alone to confirm pathogenicity. Detailed phenotypic characterization of patients to investigate phenotypic overlap is an important step to collect further evidence.⁴¹² Therefore, I aimed to confirm newly discovered NND genes and delineate the associated phenotypes by (1) the expansion of patient numbers with specific next generation NDDs using matchmaking systems and (inter)national collaborations, and

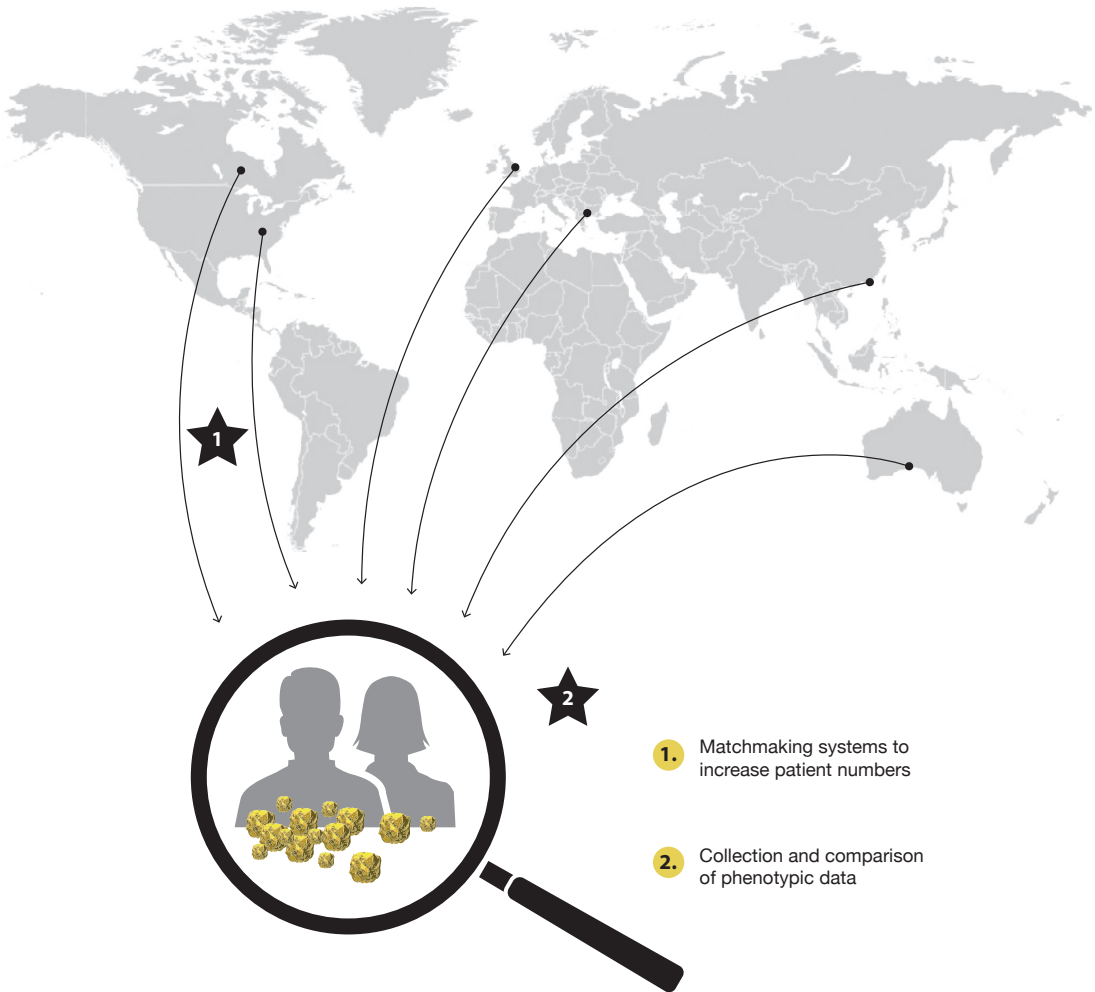


Figure 12.3: Strategies described in this thesis to confirm and delineate next generation NDDs.

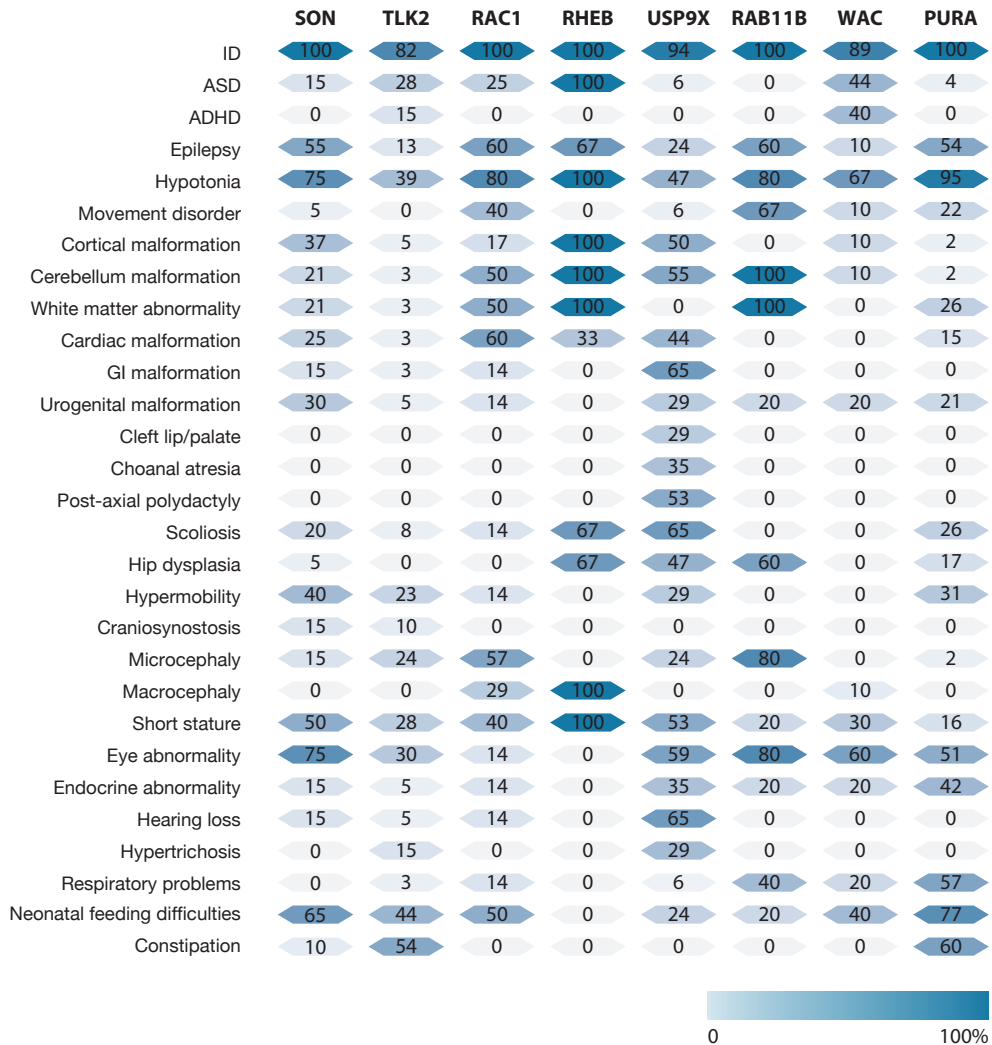


Figure 12.4: Heatmap of recurrent clinical features, representing associated phenotypes of the eight reported next generation NDDs in this thesis. Numbers indicate percentages of patients showing the feature. Abbreviations: ADHD = attention-deficit hyperactivity disorder, ASD = autism spectrum disorder, GI = gastrointestinal, ID = intellectual disability

(2) collection and comparison of detailed phenotypic data (Figure 12.3). These efforts finally resulted in the clinical characterization and confirmation of seven novel syndromes and detailed phenotypic delineation of the in 2014 discovered PURA syndrome (Figure 12.4).^{386; 387}

Increasing patient numbers with next generation NDDs by matchmaking

The importance of data sharing for rare disorders has been underscored by multiple studies, organizations and consortia.^{22; 33; 432; 433} The International Rare Diseases Research Consortium (IRDiRC), established in 2011, stated that the lack of a widely adopted data-sharing

framework and lack of common data sharing standards, currently acts as a bottleneck in the gene-discovery pipeline.²² This thesis confirms that the use of matchmaker systems and international collaborations is an effective and successful strategy: from the seven novel NDDs reported in this thesis, I was able to enlarge the cohort size in six of them (Figure 12.5). For *TLK2*, which is an ultra-rare disorder with an estimated prevalence between 0.09% and 0.24%, data-sharing finally resulted in a total number of 40 patients (**Chapter 5**). This relatively large dataset allowed confirmation of the *TLK2* gene as responsible for a non-specific form of NDD. At decreasing next generation NDD prevalence, even more extensive data-sharing will be required as shown for mutations in *RHEB*: I found no additional patients with mutations in this gene, despite worldwide data sharing and presenting of this novel NDD gene at international conferences. The three patients, including two siblings, reported in this thesis (**Chapter 7**) are to the best of my knowledge the only patients identified with *de novo* pathogenic mutations in this gene worldwide. For *RHEB*, it is likely that the extreme rarity of disease-causing mutations partly reflects a specific dominant mechanism and hence a very restricted mutational target. This hypothesis is probably also applicable to *RAB11B* and *RAC1*, which both have specific pathogenetic mechanisms with a similarly restricted repertoire of possible pathogenic mutations, and hence a very small effective mutational target.⁴³⁴ Thus, for ultra-rare NDD syndromes representing specific mutational and pathogenetic mechanisms, more extensive data-sharing will be required.

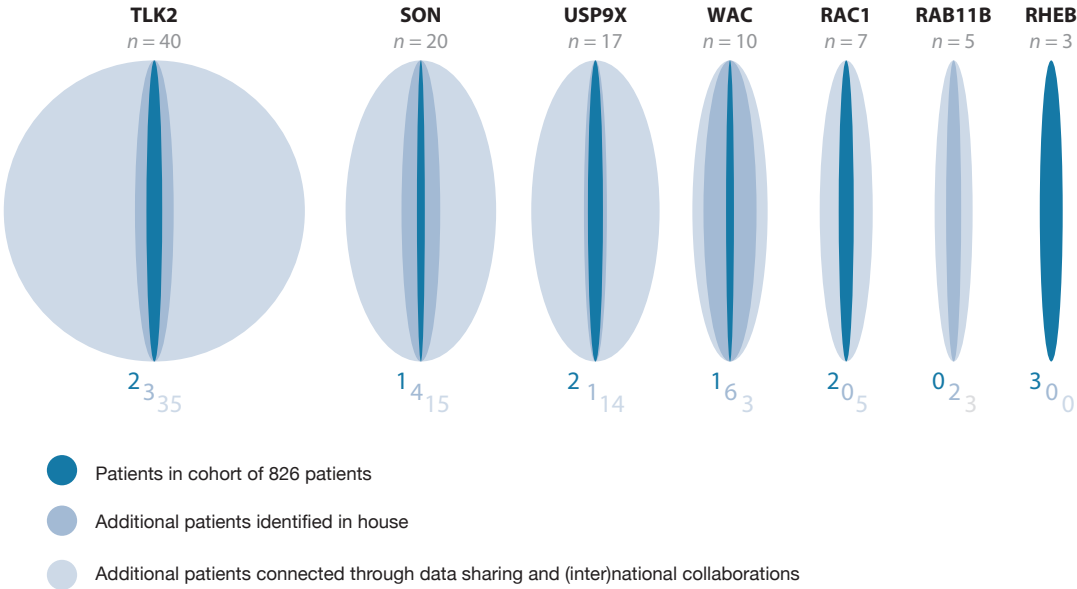


Figure 12.5: Overview of identified number of patients within the cohort of 826 patients, additional patients identified in house (n= ~4700), and additional patients identified by matchmaking and (inter)national collaborations.

Phenotypic spectrum of next generation NDDs

Most studies reporting novel NDD genes have a maximum of 15 patients included, due to limited data-availability and restricted time for patient inclusion. A few exceptions are novel NDD genes such as *DDX3X* (38 patients), *YY1* (23 patients), *STAG1* (17 patients), and *SYNGAP1* (17 patients).^{338; 435-437} But even though these numbers are relatively high for publications presenting novel syndromes, cohort sizes are still too small to have an overview about all clinical features associated with these NDDs. This makes it difficult to provide management recommendations to parents and physicians involved in the care of the diagnosed patient. Furthermore, most patients are children and only a small percentage of patients undergo WES at adult age, causing a lack of information on long term outlook, and prognostic aspects of the disorder. More longitudinal studies, and more extensive and detailed phenotypic data are needed to gain insight into the full clinical spectrum associated with the next generation NDD syndromes. For the most frequently mutated genes *ARID1B* and *ANKRD11*, there are already dozens of follow-up publications, including several studies reporting or summarizing clinical information on larger cohorts of patients.^{429; 438-441} In this thesis, I collected a large, international cohort of patients with mutations in *PURA* (**Chapter 11**), discovered as NDD gene in 2014.^{386; 387} Two years after the identification of this syndrome, the total number of patients published in literature was 22, with information distributed over five different articles. By reviewing all available clinical data and by adding a larger cohort of patients ($n = 32$), I was able to better delineate the clinical spectrum of *PURA* syndrome, providing thereby a helpful document for the counseling of *PURA* patients and their parents. One important factor in gathering clinical information on such a large cohort of patients with *PURA* mutations was the involvement of a very active parental interest group for this syndrome (*PURA* syndrome foundation at www.purasyndrome.org). To fully delineate the very rare features of the disorder, and to find answers on specific questions about features such as epilepsy (Type? Frequency? Age of onset? Treatment options?), endocrine abnormalities (Hormone levels? Clinical relevance? Pituitary gland abnormalities? Treatment needed?), or skeletal abnormalities (Severity? Age of onset? Surgery needed?), will probably require the collection of hundreds of patients and prospective studies. For ultra-rare next generation NDDs caused by mutations in genes such as *RHEB* or *RAB11B*, these numbers may remain difficult to achieve in future.

The neurodevelopmental spectrum of the next generation NDD syndromes reported in this thesis is broad (summary in figure 12.6). Discovery of next generation NDD syndromes initially started with the search for *de novo* mutations in patients with ID, but subsequent sharing of these mutations with (inter)national colleagues revealed that mutations in the same genes were also identified in patients diagnosed with other NDD subtypes. Therefore, while ID is part of all syndromes, it is not universally present. For NDDs caused by mutations in *TLK2*, *USP9X* or *WAC*, some patients had borderline ID or intelligence profiles within the normal range. In several of these patients, prevalence of other subtypes including

ASD and ADHD was reported (Figure 12.6). Genetic overlap between ID, ASD, epileptic encephalopathy and schizophrenia has already been shown in literature,¹⁰¹ and the overlap with ADHD is particularly interesting. The association of ADHD with ID is an ongoing debate: while large GWAS studies suggest a similar biological etiology, a small number of studies investigating the contribution of *de novo* mutations to non-familiar ADHD concluded that it appears to be independent from ID and ASD in the genetic pathogenesis.^{36; 442} Studies on the genetic basis of other subtypes of NDDs have been performed only sporadically. In this thesis, almost all reported patients were known to have significant language delay. Though, the presence of a formal communication disorder diagnosis was not In contrast to more mildly affected patients with *TLK2*, *USP9X* or *WAC* mutations, the neurodevelopmental phenotypes of the patients with *RAB11B* and *RHEB* mutations are the most severe end of the NDD spectrum (Figure 12.6). They are, also, the smallest cohorts in this thesis. While this likely can be explained by a restricted mutational target, it should also be considered that mutations in these genes can result in miscarriage during pregnancy, and therefore reduce the number of patients born with these syndromes. For Down syndrome and Turner syndrome it is well established that the majority of affected fetuses do not make it to term.

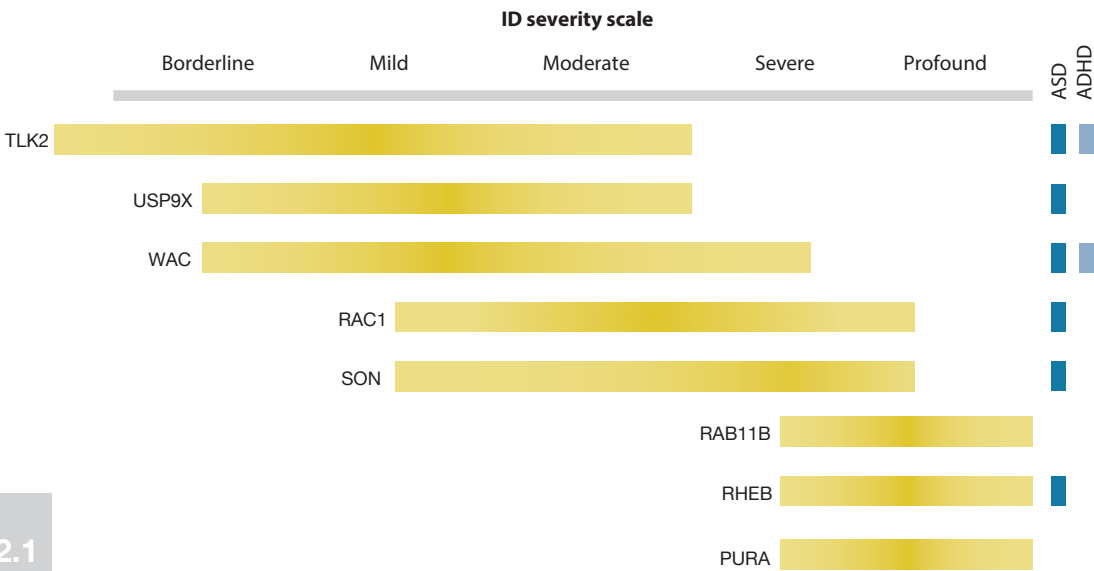


Figure 12.6: Novel next generation NDD syndromes in this thesis are associated with a variable severity of developmental delay. Disorders with milder ID levels seem to have more overlap with other NDD subtypes ASD and ADHD. Some of these patients have normal intelligence or borderline ID, but are known to have behavioral problems. Unknown is whether the syndromes are associated with NDD subtypes communication disorder, specific learning disorder and motor disorder. Clinical information about official diagnoses of these subtypes was often not available. Abbreviations: ADHD = attention deficit hyperactivity-disorder; ASD = autism spectrum disorder; ID = intellectual disability

Recently, first studies investigating deceased fetuses by WES have been published.¹¹⁰⁻¹¹² In Yates et al, mutation p.(Gly13Asp) in *HRAS* was identified in a deceased fetus with a hydrops. Exactly this mutation has been reported in a patient with Costello syndrome, a syndrome hallmarked by severe ID, macrocephaly and a variety of congenital abnormalities including heart malformations.⁴⁴³ This example shows that the neurodevelopmental spectrum of NDDs can be very broad, even if patients carry exactly the same variant, and can result in pregnancy loss in some cases, but in severely affected live births in others. Differences in timing of mutations can contribute to such phenotypic variation: postzygotic mutations (~6.5% - 7.5%)^{92; 93} can sometimes result in less severe phenotypes than germline mutations.⁷⁰

12.1.3 Next generation NDDs & biology: unraveling underlying biological mechanism

In addition to phenotypic data, functional studies on the identified mutations can be used to provide additional support for their pathogenicity.⁴¹² Furthermore, the discovery of novel genes enables the investigation of underlying biological mechanisms that may ultimately be used for the development of new therapeutics.^{433; 444} I was very fortunate to collaborate








	<div>Zebrafish</div> 	<div>Drosophila</div> 	<div>Mouse</div> 	<div>Yeast</div> 	<div>In silico modeling</div> 	<div>Patient cells</div> 	<div>Human retina cells</div> 
SON	Severe malformations spine, brain and eyes					Misspliced transcripts; downregulation critical NDD genes	
TLK2						Transcript subjected to NMD	
RAC1	Microcephaly, Reduced neuronal proliferation, Cerebellar abnormalities		Mutations genocopy DN or CA variants		Disturbed GTP activity, protein-protein interaction or protein stability		
RHEB	Macrocephaly, phenotype rescued by rapamycin		Increased soma size; aberrant neuronal migration; epilepsy; rapamycin rescue				
USP9X						Transcript subjected to NMD; protein expression	
RAB11B				Enhanced protein interaction GEF	Disturbed GTP binding or protein-protein interaction		Disturbed localization golgi and peri-centrosomal regions
WAC		Impaired learning					

Figure 12.7: Overview of insights into underlying biological mechanisms of novel identified next generation NDDs. Abbreviations: CA = constitutively active; DN = dominant-negative; NMD = nonsense-mediated decay.

with expert groups who conducted functional experiments to elucidate clues towards the underlying pathophysiological mechanism. These experiments contributed to interesting insights for all seven novel NDD genes (Figure 12.7).

Confirmation of novel NDD genes using model systems

Combining statistical evidence, recurrence of mutations, overlapping phenotypes and/or molecular findings on function collectively confirm novel NDD genes. If phenotypes are distinct and strongly overlap between patients, such as for females with *USP9X* mutations (**Chapter 9**), functional studies to prove mutation pathogenicity may be less essential. The fact that we found a localization of *USP9X* to the primary cilium is interesting information to further understand disease mechanisms, but was not essential evidence to confirm *USP9X* as novel NDD gene in females. Similarly, finding large numbers of patients with mutations in the same gene reduces the need for extensive functional validation. The best example in this thesis is *TLK2* (**Chapter 5**). With the identification of 40 patients with *de novo* mutation in this gene, *TLK2* could be added to the list of NDD genes, beyond reasonable doubt. However, the majority of novel next generation syndromes have less distinct phenotypes and much smaller patient cohorts. For these genes, functional confirmation is more important, in particular if missense mutations predominate. In general, interpretation of missense mutations, comprising more than 60% of *de novo* mutations identified by WES, is more challenging than truncating mutations such as nonsense, frameshift or splice site mutations.⁴¹² Missense mutations within the same gene can result in different outcomes, depending on the localization of the mutation and the potential to modify protein function, as we showed for mutations in *RAC1* (**Chapter 6**).^{412; 445} Given that functional work, as performed for the with missense associated genes *RAB11B*, *RAC1* and *RHEB*, is time-consuming, it may be difficult to provide functional evidence for every mutation in a novel next generation NDD gene. Nevertheless, there are initiatives aiming to do just this. For example, the Undiagnosed Disease Network (UDN) in the United-States has set up an extensive workflow for patients with undiagnosed rare disease, including NDDs. In this workflow, clinical characteristics of patients without a mutation in a known NDD gene are sent to researchers, who compare variants of unknown pathogenicity with existing information on the gene/variant in model systems. For interesting variants in candidate NDD genes, zebrafish and *Drosophila* models are developed to confirm variants and novel NDD genes.⁴⁴⁴ Similarly, large-scale government funded projects such as Care4Rare in Canada, and more recently EU-funded projects within the European Reference Networks (SOLVE-RD), are setting up broker services to link-up genomics researchers with laboratories performing functional experiments with the aim of speeding up and promoting functional validation of novel disease genes and their mutations. Confirmation of novel NDD genes would make headway if comparable workflows are available for all patients with undiagnosed NDDs. But since most countries have no sponsors such as the National Institutes of Health (NIH; sponsor of UDN), the initiation and maintenance of such large structural collaborations and

infrastructures is difficult. Techniques accelerating the ability to measure and interpret the functional consequences of identified variants, such as saturation genome editing, can be thereby helpful.⁴⁴⁶

Further insights into underlying biological mechanisms

Biological findings reported in this thesis can act as model for other genes, proteins and pathways. As an example, loss-of-function mutations in *SON*, encoding a protein with an important role in gene transcription and RNA splicing,^{128; 133-136} disturb the function of multiple other cortical developmental and metabolic proteins by accumulation of mis-spliced transcripts (**Chapter 4**). Pur-Alpha or PURA (**Chapter 11**), is involved in similar cellular processes as SON: PURA is known to play a role in DNA replication and transcription, but also in RNA splicing.^{392; 393; 447} Recently, a case report was published describing a patient with a pathogenic *PURA* mutation.⁴⁴⁸ In this patient, hypoglycorrachia was observed in addition to the typical PURA syndrome features. This clinical feature is a rare metabolic finding often caused by *SLC2A1* point mutations or deletions.⁴⁴⁹⁻⁴⁵¹ The authors indeed demonstrated significant downregulation of *SLC2A1* in patient cells, likely explaining the patient's unusual phenotype. But contrary to expectations, no point mutation in or deletion of *SLC2A1* was identified. Therefore, the authors proposed that *SLC2A1* might be a target of PURA.⁴⁴⁸ Interestingly, in at least one other PURA patient, borderline-low CSF glucose levels have been measured, and again the patient tested negative for a *SLC2A1* mutation [personal observation]. Taken into account this observation, and the reported function(s) of PURA, including essential cellular processes such as DNA replication, transcription and RNA splicing, it would be interesting to perform systematic transcriptome studies in patient-derived cells, using the SON study as a model, to investigate whether *PURA* mutations lead to downregulation of *SLC2A1* and possible other neurodevelopmental and/or metabolic genes.

In this thesis, missense mutations were identified in five genes: *TLK2*, *RAC1*, *RHEB*, *USP9X* and *RAB11B*. In three of these, *RAC1*, *RHEB* and *RAB11B*, *in silico* modeling and functional experiments showed that mutations likely act through gain-of-function or dominant-negative mechanisms (**Chapters 6-8**). Interestingly, all three genes belong to the same human RAS superfamily of small GTPases, which all are activated by a similar biochemical activity: GTP binding and hydrolysis.¹⁷⁶ If bound to GTP, the proteins display a binding surface, formed by the two switch regions, that have high affinity for downstream effectors. In contrast, if GTP is hydrolyzed, the protein is mainly inactive, since effector proteins are released due to reduced affinity, thereby attenuating downstream signaling.¹⁷⁶ Both the GTP binding and switch regions are highly conserved among RAS proteins and different species.¹⁷⁶ The missense mutations in *RAC1*, *RHEB* and *RAB11B* all occurred within, or close to, these regions (Figure 12.8A). In literature, *de novo* missense mutations in *RAB11A*, homologue of *RAB11B*, were recently identified in two patients with epilepsy and moderate-severe ID.⁴²

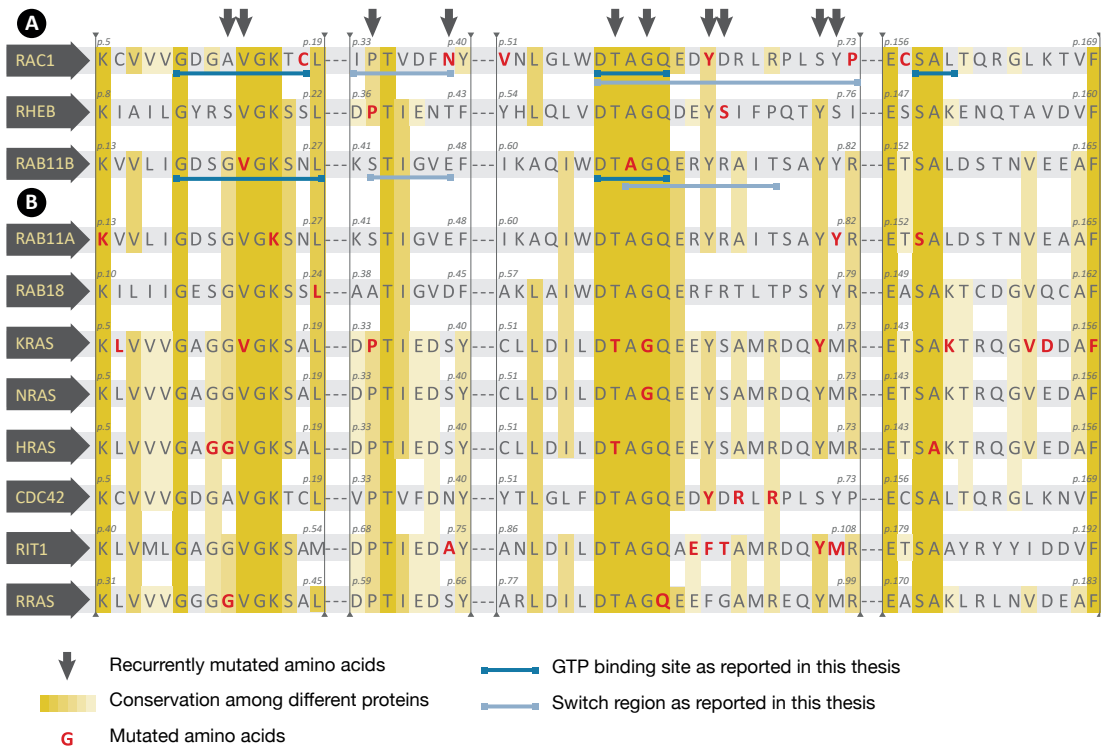


Figure 12.8: Alignment of parts of human RAS protein members known to be associated with NDDs. Reported affected amino acids are shown in red, conserved amino acids in more than five proteins have a yellow background. Blue lines indicate either a GTP binding site (dark blue) or Switch region (light blue), as reported in Chapter 6 (RAC1) and Chapter 8 (RAB11B). The small arrows pointing the columns indicate amino acids that are mutated in different proteins ('recurrent mutations') and show the clustering of mutations between different proteins in often highly conserved amino acids. Other observed mutations affect amino acids that are often very close to the amino acids with recurrent mutations. (A) Alignment of the three RAS protein members identified in this thesis as novel NDD genes. (B) Alignment of eight RAS proteins reported in literature as NDD genes.

Meta-analysis revealed that three additional *de novo* missense mutations were reported by the DDD study. Also another RAB protein, RAB18 (RAB subfamily), has been associated with NDD⁴⁵², as well as KRAS, HRAS, RRAS, NRAS, and RIT1 (RAS subfamily)^{423; 443; 453-455}, and the recently published CDC42 protein (RHO subfamily). Similar to *RHEB*, *RAC1* and *RAB11B*, the majority of reported missense mutations in *RAB11A*, *RAB18*, *NRAS*, *KRAS*, *RRAS*, *HRAS*, *RIT1* and *CDC42* are located within such GTP binding and switch regions (Figure 12.8B).^{443; 452; 454; 456-466} Interestingly, when aligning the mutations in these genes observed in NDD patients, mutations seem to cluster: reported pathogenic missense mutations in genes encoding RAS proteins often affect the same, highly conserved regions, which represent a limited number of amino acids. Even more striking is the fact that some mutations affect exactly the same amino acid (for example: the proline in both *RHEB* and *KRAS*; the tyrosine in both *RAC1* and

CDC42; the valine in both *RAB11B* and *KRAS*; Figures 12.8A and B). Recently, a study has been published suggesting that the use of ‘meta-domains’ can improve the interpretation of genetic variation.

Authors found that genetic tolerance is consistent across protein domain homologues, and that patterns of genetic (in)tolerance mimic patterns of evolutionary conservation.⁴⁶⁷ Other studies reported that spatial clustering of mutations can lead to gene identification.^{468; 469} In one of these studies, both *RAB11A* and *RAB11B* were found to have significant clustering of missense mutations at protein level relative to population controls.⁴⁶⁹ It would be interesting to investigate whether these statistical models can be combined and expanded: mutations in other RAS proteins could potentially be discovered by combining clustering and meta-domains, since mutations in genes encoding RAS proteins seem to cluster.

It is not unlikely that subsets of pathogenic mutations in different RAS proteins disrupt RAS protein mechanisms in a similar way. Some will disturb the nucleotide binding pocket, resulting in either reduced or increased GTP binding. For example, missense mutations in *RAB11B* likely impaired RAB11B GTP binding (**Chapter 8**), given the mis-localization within the cell that mimicked the inactive state of the protein, and the abnormal high binding affinity towards a GEF. A mutation in another gene, *HRAS*, is only two amino acids away from mutated amino acid 22 in *RAB11B* (Figures 12.8A and B), and has been reported to have decreased GTPase activity.⁴⁷⁰ On the other hand, mutations disturbing exactly the same amino acid (tyrosine) in both *RAC1* and *CDC42*, are both thought to have the opposite effect: in **Chapter 6** I showed that *RAC1* mutation p.Tyr64Asp had a constitutively active effect, and Martinelli et al.⁴⁶⁶ reported an increased amount of active, GTP-bound protein for *CDC42* mutation p.Tyr64Cys. Mutations in *HRAS* causing such a hyperactive state of the protein have been reported to act as ‘selfish mutations’: mutations that cause a higher self-renewal than surrounding wild-type cells and expand clonally in testis with higher paternal age.^{471; 472} The mutations lead to relative enrichment of mutant sperm over time, resulting in a higher risk for the father to pass the mutation to one or more children. It would be interesting to investigate whether activating mutations in other RAS proteins similarly have a paternal age effect. Inclusion of information on the age of the parents at birth of each child in articles will be therefore important in future studies.⁴⁷¹

In contrast to rare disorders caused by mutations with major effect sizes (monogenic), common diseases such as hypertension, diabetes, obesity, and scoliosis are mostly due to multiple genetic factors each contributing with smaller effects. Common diseases often include a small subset of individuals with a monogenic form of the disease.⁴³³ Although these variants explain only a small fraction, they make a significant contribution to our understanding of biological mechanisms underlying common disease and, ultimately, the development of therapy.^{433; 473; 474} It is estimated that approximately 20% of genes implicated

in monogenic forms of common disorders, also contain a variant responsible for a GWAS signal that achieves genome-wide significance for a complex trait.⁴³³ In this thesis, I showed that rare, monogenic causes of syndromic NDDs are not only relevant to find genetic causes of complex traits, but can also be used to find associations between gene sets or pathways and normal variation in the population. In **Chapter 7**, I reported that NDD syndromes caused by disruptive mutations in genes related to MTOR, were strongly associated with megalencephaly in NDD patients. In line with this, common variation in these genes was found to be significantly associated with intracranial volume in the population, thereby showing that disruptive mutations causing rare NDD syndromes can lead to an extreme phenotype beyond a continuing phenotypic spectrum in the population that can be explained by common variation in a similar gene set. This finding raises the question if other phenotypic extremes in NDD syndromes can be associated with other forms of phenotypic variation in the population. For example, NDD syndromes caused by mutations in genes involved in the bromodomain family III, such as *EP300*, *CREBBP*, *BRWD3* and *PHIP*, frequently co-occur with obesity.⁴⁷⁵⁻⁴⁷⁷ One might investigate whether common variation in genes related to this bromodomain family III are associated with body weight in population, a characteristic known to have a high heritability and to have genes involved in the central nervous system as critical regulators.⁴⁷⁸⁻⁴⁸¹

12.2 Neurodevelopmental disorders: what's next?

12.2.1 Diagnostic and research strategies for undiagnosed NDD patients

Larger patient cohorts, increasing power (Figure 12.9)

Approximately 50% of patients referred for next generation sequencing remain undiagnosed at molecular level.²² Also in the cohort described in this thesis, a significant percentage of patients received no molecular diagnosis. For these patients, the ‘diagnostic odyssey’ is not yet finished. The importance of a continued search for a diagnosis in these patients is supported by large (international) consortia, such as the IRDiRC, the UDN, and the European Reference Network (ERN) ITHACA:^{22; 432; 482} “Exact disease recognition, an element of the concept of precision in medicine, requires new infrastructure that spans geography and institutional boundaries”.⁴³² An universally accepted and implemented system of international

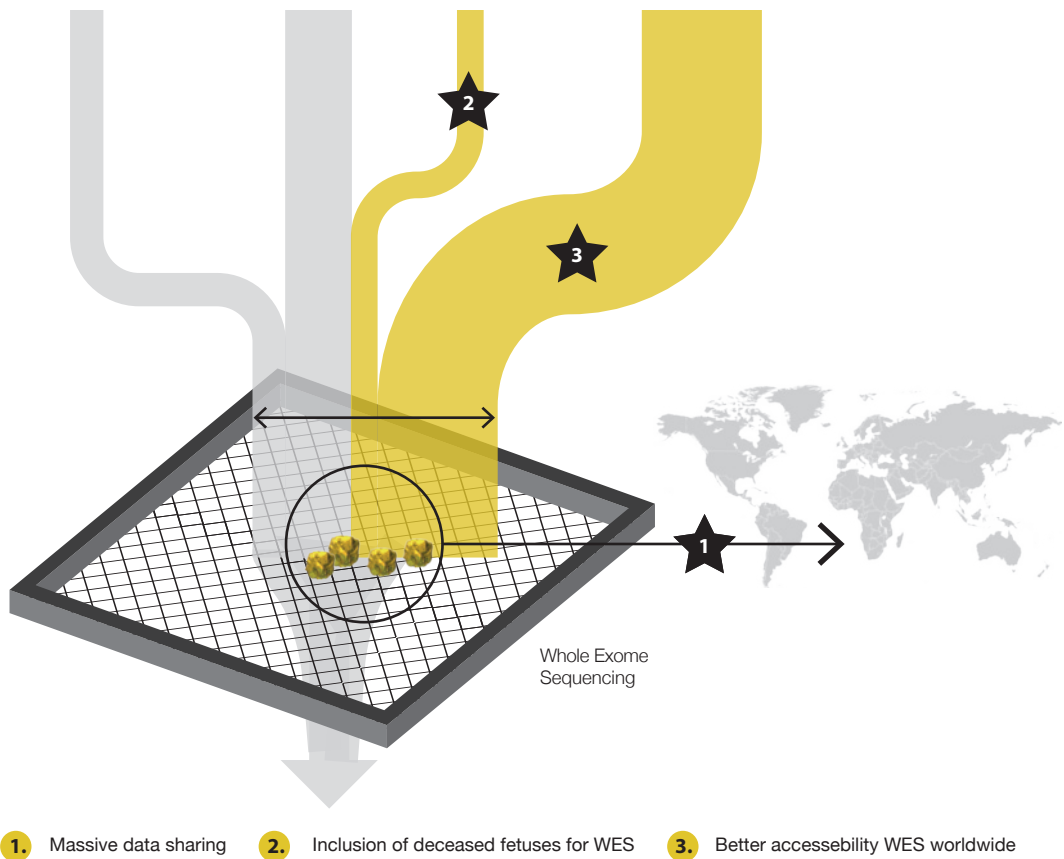


Figure 12.9: Strategies to enlarge patient cohorts, thereby increasing the power for the discovery of novel NDDs.

large-scale data-sharing, in this thesis shown to be successful on a small scale, will be a major challenge to achieve over the next years. But if established, it will lead to enlarged cohort sizes, thereby further increasing the power to find recurrence of mutations and to find novel NDD genes. The contribution of parents to data collection and datasharing should not be underestimated: where clinicians and researchers are more often reluctant in sharing data, parents can be much more proactive, partly because data sharing is sometimes the only possibility to receive a diagnosis for their child.

To further enlarge patient cohorts and the corresponding WES datasets, it will be important to include and test more mildly affected patients, as shown in this thesis. In addition, it would be interesting to systematically collect data of deceased fetuses, who may sometimes represent the most severe end of the NDD phenotypic spectrum. For example, in *Vora et al.*, a *de novo* nonsense mutation in *MAP4K4* was identified in a fetus with multiple congenital abnormalities, which often co-occur with NDDs. Authors concluded that this mutation could not be classified as disease causing, since it was not associated with disease before.¹¹² But, inspection of the international *de novo* database²⁵² shows that two other truncating mutations in *MAP4K4* have been reported in literature in patients with NDD.³³ Additionally, four *de novo* missense mutations were present: three in NDD patients and one in a patient with a syndromic congenital heart defect.^{33; 483} In ExAC, *MAP4K4* is classified as gene highly intolerant for both loss-of-function and missense variation in the population (z-score of 4.01 and pLI score of 1.00). Based on this information, it is not unlikely that mutations in *MAP4K4* indeed cause a syndromic NDD, of which intra-uterine death is the end of the phenotypic spectrum. I expect that increasing the NDD cohort size by adding WES data of deceased fetuses to large NDD WES datasets, will facilitate the discovery of more NDDs, in particular those associated with congenital malformations.

Finally, wider availability of WES will increase the collective power of the community to identify novel NDDs. An illustrative example can be observed in the origin of patients with PURA syndrome: ten patients included in the study reported in **Chapter 11** were patients tested in the Netherlands, but no patients were identified in European countries such as Belgium and Germany. An important difference is that in the Netherlands the costs of WES are covered by the national health insurance, which is not the case in many other European countries. For this reason, most NDD patients in these countries remain untested. This currently also applies to the majority of (large) non-Western countries such as China.⁴⁸⁴ Accessibility of WES for all NDD patients in Europe, an important goal of the ERN ITHACA,⁴⁸² but also worldwide as envisioned by IRDiRC, will contribute on large scale to the enlargement of patient cohorts and datasets, thereby raising power to discover novel rare next generation NDDs.

Improved interpretation of WES data (Figure 12.10)

Recent re-evaluation of WES data already has shown that the diagnostic yield of WES was increased up to 15.4% within a few years.^{100; 103-108} In the future, a further increase can be obtained by improved interpretation of WES data, leading to more diagnoses and discovery of more next generation NDDs. Besides the four suggestions (Figure 12.10A-D) mentioned in the first part of the discussion, other strategies can be considered.

Incorporation of standardized clinical data in the diagnostic process (Figure 12.10E)

Incorporation of clinical information on the patient, preferably in HPO terms⁴²⁴, in the diagnostic process will lead to better interpretation of WES data. In this thesis I reported that 26% of patients with macrocephaly had a *de novo* mutation in a gene related to mTOR. This is a high percentage of patients with a distinct overlapping phenotype harboring mutations in a small subset of genes. In diagnostics, this information can be used for better interpretation of WES data. If a patient has macrocephaly, and this information is available in the diagnostic pipeline, special attention could be paid to variants within the set of mTOR-related genes. This example illustrates the need for a direct link between clinical and diagnostic settings.

Implementation of improved statistical and computational models and expansion of control databases (Figures 12.10F and G)

Computational models currently develop fast and will further improve data interpretation.²² Prioritizing genetic variants by tissue-specific prediction models⁴⁸⁵, assessing the likelihood for a gene to harbor dominant changes by the machine-learning approach DOMINO⁴⁸⁶, and prioritizing synonymous SNVs based on a computational algorithm⁴⁸⁷ are just a few examples of recently published models developed as useful components to NGS pipelines. Although such models should be carefully reviewed before implemented, it will probably lead to increased possibilities to filter out the (possibly) pathogenic variants, which previously remained unrecognized. The use of appropriate databases containing genomic data of healthy controls is thereby essential. Databases such as ExAC⁷¹ and GnomAD provide these data, but one should bear in mind that variants in these databases are often not validated. Additionally, while these databases focus on 'healthy' controls, they may still contain a limited amount of data from mildly affected individuals. Variant interpretation is also more challenging for patients from non-European/North-American ancestry: due to relatively small numbers of patients from these populations in the control databases, it is harder to separate pathogenic variants from rare benign background genetic variation.^{65; 104} Initiatives from organizations such as the Global Alliance for Genomics and Health have been established to include more samples from such underrepresented populations.⁴⁸⁸

More attention for X-linked variants in females (Figure 12.10H)

Current research mainly focuses on the identification and analysis of *de novo* mutations, which has been very successful so far. However, pathogenicity of *de novo* variants on the

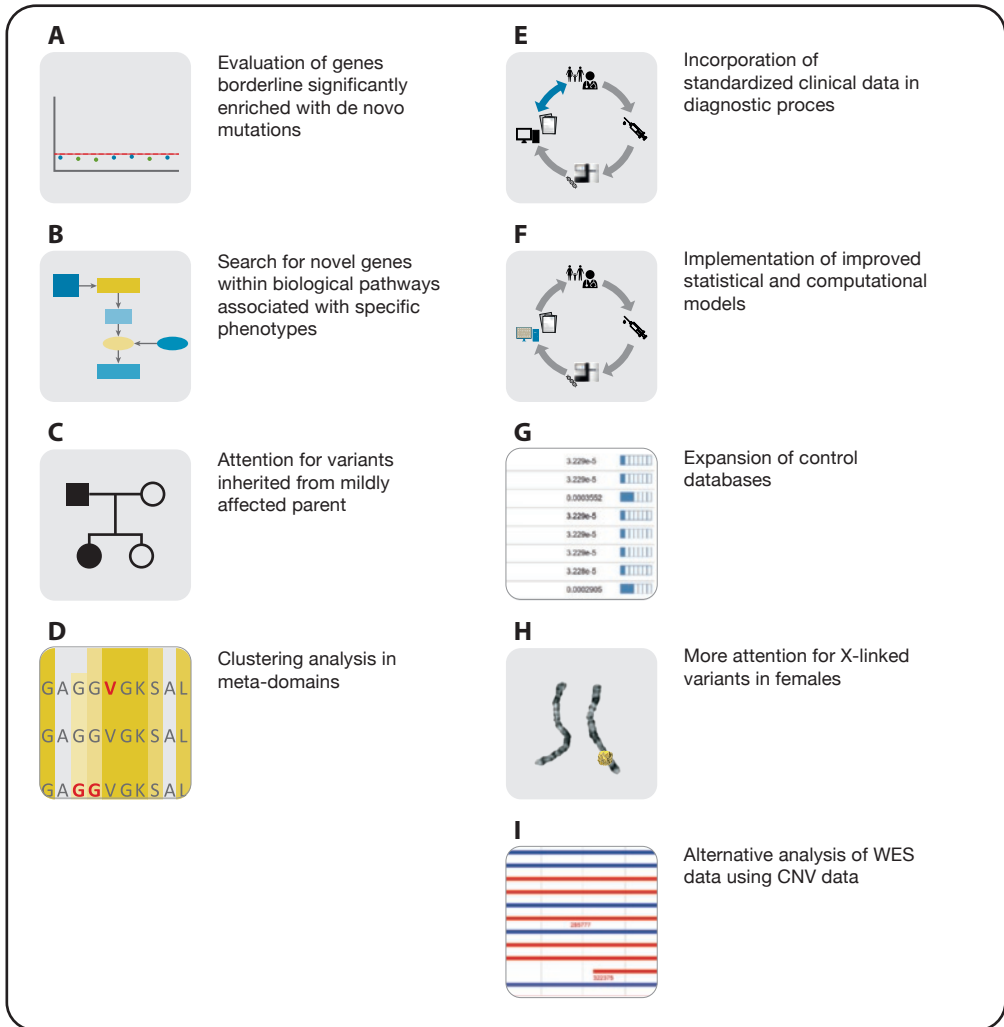


Figure 12.10: Summary of strategies to improve interpretation of WES data, thereby enabling the discovery of novel NDDs. Strategies A – D were mentioned in Chapter 12.1, whereas strategies E – I are further explained below.

X-chromosome in females is sometimes not appreciated. Historically, X-linked disorders occurred mainly in recessive form in males, with unaffected or mildly affected heterozygous females transmitting the mutation to their son(s). Since females have two X-chromosomes of which one can be preferably inactivated, they are sometimes thought to be protected against X-linked disorders. Mutations in X-linked *MECP2*, causing Rett syndrome in females,⁴⁸⁹ is a clear exception to this rule. The identification of *de novo* variants in *USP9X* in females (**Chapter 9**), as well as some other next generation genes such as *DDX3X*³³⁸, *SMC1A*⁴⁹⁰, *HUWE1*⁴⁹¹, *KIAA2022*⁴⁹² constitute an expanding list of dominantly inherited NDDs caused by *de novo* mutations on the X-chromosome. Also the DDD study reported a genome-wide significance excess of *de novo* mutations in females in a few genes, including *USP9X*.³³ Interestingly, the majority of newly identified genes are known to (partially) escape X-inactivation. These escaping genes appear to cluster in a few Topological Associating Domains (TADs).^{330; 493; 494}

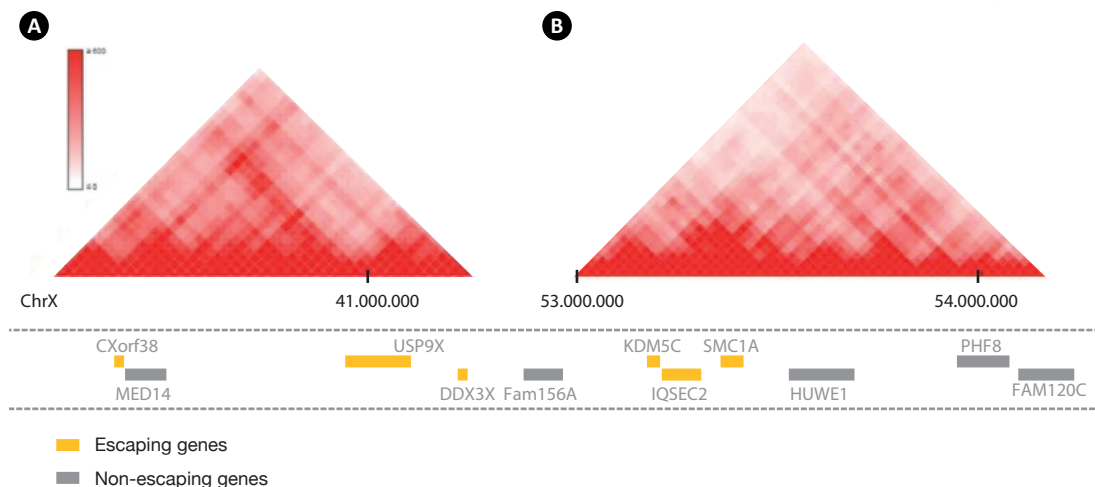


Figure 12.11: NDD genes *USP9X* and *DDX3X*, as well as *SMC1A* and (candidate) NDD genes *KDM5C* and *IQSEC2* are known to cluster into the same Topological Associating Domain and all (partially) escape X-inactivation.

Indeed, *USP9X* and *DDX3X* are located directly next to each other in a TAD, whereas *SCM1A* is located in another TAD together with *KDM5C* and *IQSEC2*, both marked as (candidate) NDD genes in females (Figure 12.11).⁴⁹⁵⁻⁴⁹⁸ With more research on and improved interpretation of X-linked variants in females, particularly in genes clustering in these TADs associated with escaping genes, more X-linked genes could be discovered.

Alternative analyses of WES data using CNV data (Figure 12.10I).

Currently, a combination of SNV and CNV analysis in WES data is the most cost-effective

genetic test with the highest diagnostic yield.^{49; 53} Previously, micro-arrays enabling the detection of CNVs were the first-tier test for patients with NDDs or congenital abnormalities.^{499; 500} In 2014, Coe et al. showed that micro-array data could be combined with WES data to successfully discover new syndromes and genes involved in NDDs.⁵⁰¹ In this thesis, some of the discovered genes were implicated in genomic regions associated with structural variants causing NDDs: *SON* at 21q22.11; *WAC* at 10p11.23 and *PURA* at 5q31.2q31.3. Nevertheless, studies structurally investigating combined CNV- and SNV datasets are lagging behind on studies focusing on WES data only. I hypothesize that interpretation of WES data can be improved by (1) the search for *de novo* variants in genes associated with single-gene deletions, (2) the combination of WES and array data to systematically evaluate regions and genes associated with specific phenotypes, and (3) the comparison of large CNV- and SNV datasets, including CNVs identified in WES data. Supporting evidence for this hypothesis comes from literature. First, several studies focusing on the analysis of array-data mentioned the identification of deletions encompassing only one gene. In these studies, single-gene deletions were found in genes such as *TRIO* and *TRIP12*.^{19; 502} *TRIO* has since been confirmed to be a true NDD gene by the identification of *de novo* point mutations, similar to *TRIP12*, in this thesis reported as candidate NDD gene and confirmed by others.^{193; 194; 405} Secondly, a high-resolution map of pathogenic phenotypes associated with their respective genomic regions has been published recently.⁵⁰³ Such regions associated with specific phenotypes can be used for the identification of novel NDD genes and syndromes with specific phenotypes. As an example, Conte et al. used overlapping CNVs associated with the specific phenotype 'oral facial cleft' to define candidate genes for this phenotype.⁵⁰⁴ These candidate genes would be an ideal starting point to search for point mutations in patients with a comparable phenotype. For NDDs, specific phenotypes such as megalencephaly, cortical brain malformations, or epilepsy, that often co-occur with NDD, can be used for a similar analysis. Finally, literature reporting on successful CNV analysis on WES data has been published.^{102; 505} Results of these analyses should be added to and combined with datasets containing micro-array data and WES data on SNVs.

Implementation of Whole Genome Sequencing in diagnostics (Figure 12.12)

Although WES has proven successful for the identification of NDD disorders, it also has limitations. In the first years of WES in diagnostics, fluctuation in coverage of any of the three samples from the trio analysis could influence the reliable detection of *de novo* variants.⁵⁰⁶ Currently, in particular the first exons and single-exon genes remain difficult to capture due to high GC content, and mutations here can be easily missed.^{106; 507} In addition, CGG repeat expansions, such as in Fragile X syndrome, cannot be mapped because of the high repeat content of the expansion.⁵³ It also has been mentioned that existing exome sequencing kits have limitations, since knowledge of all protein-coding exons in the genome is still incomplete.⁹⁷ Finally, mutations located in non-coding regions such as regulatory- or deep

intronic regions and complex structural variants cannot be detected by WES.⁵⁵ Whole genome sequencing (WGS) and long-read sequencing technologies have been shown to (partially) overcome these limitations.⁵⁰⁸ Recent studies reported a diagnostic yield of WGS between 41% and 77% in patients with undiagnosed NDDs, which were negatively tested with WES.⁴⁶⁻⁴⁸ Positive results included pathogenic non-exonic sequence variants and structural variants which were not detected by WES. Although these percentages are promising, in particular for patients without any identified *de novo* mutation by WES, there are disadvantages delaying the implementation in diagnostic settings. High costs for WGS (WES costs are 30-40% of WGS costs), which are not covered by health insurances yet, is the most important disadvantage.^{22; 33} Lowering costs will pave the way to implement WGS as diagnostic test for undiagnosed NDD patients, thereby increasing diagnostic possibilities for these patients.

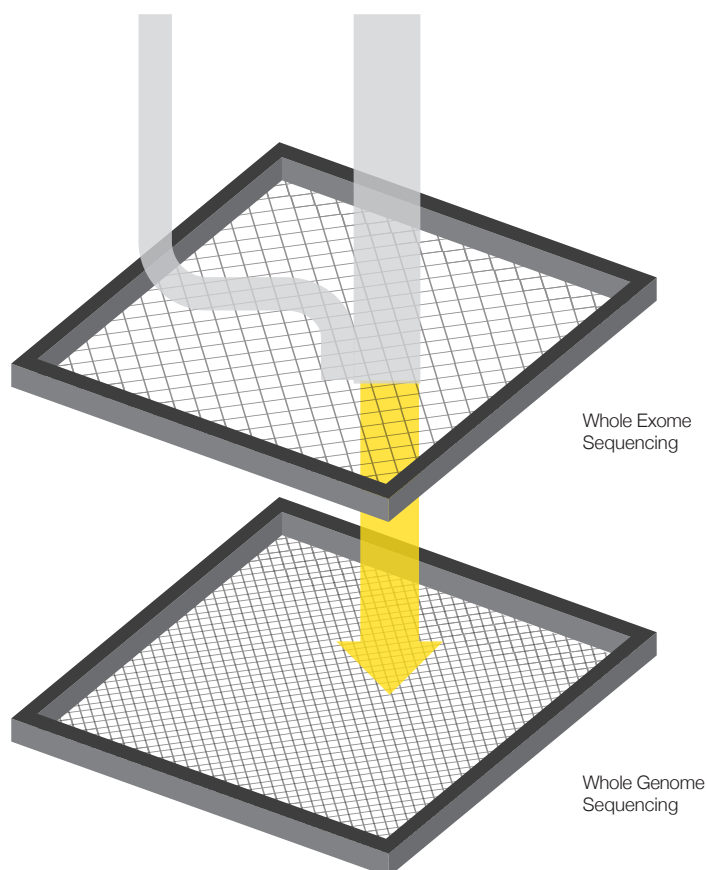


Figure 12.12: Strategy to discovery novel NDDs by implementation of WGS.

12.2.2 Improved patient care for next generation NDDs

The massive identification of next generation NDD syndromes already has a significant impact on patient care. The next challenge is to translate the discovered NDDs in clinic and to improve care for patients and their parents. Only then, the ultimate goal of gene identification can be achieved: precision medicine for the NDD population.²² Projects such as gene-specific leaflets from Unique (www.rarechromo.org), Dutch national expertisecenters for rare diseases including rare NDDs (www.nfu.nl), and virtual multidisciplinary teams across EU centers, accessible by Telehealth technology (www.ernithaca.org) are just a few examples of initiatives attempting to improve care for patients with rare disorders, including next generation NDDs.

Parent empowerment: the tale of the PURA Foundation

Soon after the discovery of *de novo* mutations in *PURA* as cause of a novel NDD by the end of 2014,^{386; 387} I became as researcher and physician involved in this syndrome. In the meta-analysis study (**Chapter 3**) we were working on at that time, we found four patients with a *de novo* mutation in *PURA*. Contact with the Southampton group, who published one of the first two reports,^{386; 387} about our patients was easily made. Soon after our first contact, I was introduced to a small, but enthusiastic group of parents of recently discovered 'PURA children'. They had ambitious plans to set up a foundation for the support and education of patients and their families, and to fund research related to PURA syndrome. During the long process of official registration of the foundation, they created a website (www.purasyndrome.org) and initiated a parental Facebook group. An official management board was formed and clinicians from different countries were asked to form a medical advisory team. After being officially registered as foundation, fund raising events could be started. Because of the growing numbers of patients identified with PURA syndrome, the plan was raised to organize the first official PURA Syndrome Foundation conference. In June 2016, only 1.5 years after the discovery of the syndrome, the first international conference took place in London, UK. Besides patients, parents, and clinicians, researchers from different continents with interest in the PURA protein were present. During the conference, a research plan was drafted, finally resulting in the first PURA research team publication, in this thesis presented in **Chapter 11**. Currently, more functional studies are ongoing (www.purasyndrome.org/rdd2018), and elaboration of plans to start a global patient registry are in advanced stage, funded by the PURA Foundation. The establishment of a research team and medical team in different countries, the registration as an official foundation, the organization of two international conferences, funding of research and a global patient registry: parents of PURA children showed that is all possible within 3 years after discovery of a novel NDD gene. The PURA Foundation is a unique example that parents can achieve what is impossible for physicians and researchers, due to a lack of research time and money.

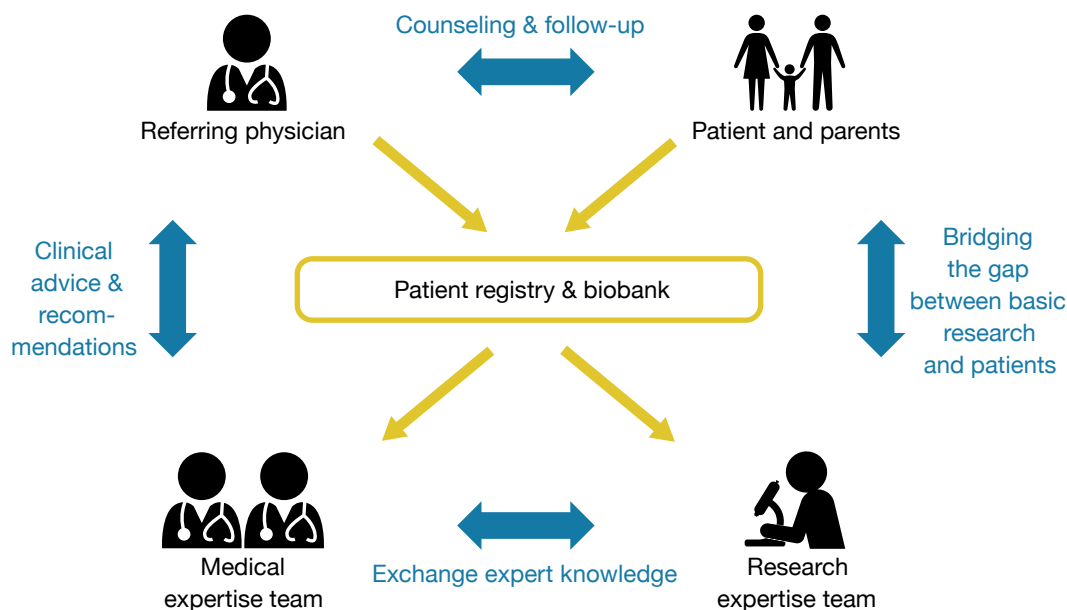


Figure 12.13: Proposed organizational structure for the care of patients with next generation NDDs.

The PURA story as an example for other next generation NDDs

Although I realize that it often takes much more time to achieve what parents of PURA children did within three years, I believe that the organizational structure of the PURA Foundation, summarized in Figure 12.13, ideally can function as an example for other rare next generation NDDs. By involving parents in the organization and bundling of knowledge in expertise teams (both clinical and research), a lot can be achieved with the use of minimal time and money, which is often the limiting factor in follow-up research. A regular conference where patients, parents, clinicians and researchers meet, will bridges the gap between different disciplines and will inspire everyone involved. Important to note is that expertise teams should contain members from different countries, since patients with the rare NDDs are diagnosed worldwide. Not only referring clinicians, but also parents should have a role in the collection of clinical data, to keep registered clinical information up to date.





A

Appendices

Reference list

1. (2013). American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th edition). (Washington DC).
2. Organisation., W.H. (2016). ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic Guidelines. (Geneva).
3. (AAIDD), A.A.f.I.a.D.D. (2010). https://aaidd.org/intellectual-disability/definition#.WglsX61x_BI. In. (
4. Maulik, P.K., Mascarenhas, M.N., Mathers, C.D., Dua, T., and Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res Dev Disabil* 32, 419-436.
5. Brugha, T.S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., Bebbington, P., Jenkins, R., and Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Arch Gen Psychiatry* 68, 459-465.
6. Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., and Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet* 368, 210-215.
7. Autism, Developmental Disabilities Monitoring Network Surveillance Year Principal, I., Centers for Disease, C., and Prevention. (2012). Prevalence of autism spectrum disorders--Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. *MMWR Surveill Summ* 61, 1-19.
8. Polanczyk, G., de Lima, M.S., Horta, B.L., Biederman, J., and Rohde, L.A. (2007). The worldwide prevalence of ADHD: a systematic review and metaregression analysis. *Am J Psychiatry* 164, 942-948.
9. Simon, V., Czobor, P., Balint, S., Meszaros, A., and Bitter, I. (2009). Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. *Br J Psychiatry* 194, 204-211.
10. Thomas, R., Sanders, S., Doust, J., Beller, E., and Glasziou, P. (2015). Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics* 135, e994-1001.
11. Altarac, M., and Saroha, E. (2007). Lifetime prevalence of learning disability among US children. *Pediatrics* 119 Suppl 1, S77-83.
12. Erskine, H.E., Ferrari, A.J., Nelson, P., Polanczyk, G.V., Flaxman, A.D., Vos, T., Whiteford, H.A., and Scott, J.G. (2013). Epidemiological modelling of attention-deficit/hyperactivity disorder and conduct disorder for the Global Burden of Disease Study 2010. *J Child Psychol Psychiatry* 54, 1263-1274.
13. Law, J., Boyle, J., Harris, F., Harkness, A., and Nye, C. (2000). Prevalence and natural history of primary speech and language delay: findings from a systematic review of the literature. *Int J Lang Commun Disord* 35, 165-188.
14. Lingam, R., Hunt, L., Golding, J., Jongmans, M., and Emond, A. (2009). Prevalence of developmental coordination disorder using the DSM-IV at 7 years of age: a UK population-based study. *Pediatrics* 123, e693-700.
15. Scharf, J.M., Miller, L.L., Gauvin, C.A., Alabiso, J., Mathews, C.A., and Ben-Shlomo, Y. (2015). Population prevalence of Tourette syndrome: a systematic review and meta-analysis. *Mov Disord* 30, 221-228.
16. Robertson, M.M. (2008). The prevalence and epidemiology of Gilles de la Tourette syndrome. Part 1: the epidemiological and prevalence studies. *J Psychosom Res* 65, 461-472.
17. Centers for Disease, C., and Prevention. (2009). Prevalence of diagnosed Tourette syndrome in persons aged 6-17 years - United States, 2007. *MMWR Morb Mortal Wkly Rep* 58, 581-585.

18. The Deciphering Developmental Disorders, S., and The Deciphering Developmental Disorders, S. (2014). Large-scale discovery of novel genetic causes of developmental disorders. *Nat New Biol*.
19. Vulto-van Silfhout, A.T., Hehir-Kwa, J.Y., van Bon, B.W., Schuurs-Hoeijmakers, J.H., Meader, S., Hellebrekers, C.J., Thoonen, I.J., de Brouwer, A.P., Brunner, H.G., Webber, C., et al. (2013). Clinical significance of de novo and inherited copy-number variation. *Hum Mutat* 34, 1679-1687.
20. Sheridan, E., Wright, J., Small, N., Corry, P.C., Oddie, S., Whibley, C., Petherick, E.S., Malik, T., Pawson, N., McK-inney, P.A., et al. (2013). Risk factors for congenital anomaly in a multiethnic birth cohort: an analysis of the Born in Bradford study. *Lancet* 382, 1350-1359.
21. Ropers, H.H. (2010). Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 11, 161-187.
22. Boycott, K.M., Rath, A., Chong, J.X., Hartley, T., Alkuraya, F.S., Baynam, G., Brookes, A.J., Brudno, M., Carracedo, A., den Dunnen, J.T., et al. (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *Am J Hum Genet* 100, 695-705.
23. Yamada, A., Kato, M., Suzuki, M., Suzuki, M., Watanabe, N., Akechi, T., and Furukawa, T.A. (2012). Quality of life of parents raising children with pervasive developmental disorders. *BMC Psychiatry* 12, 119.
24. Jonsson, U., Alaie, I., Lofgren Willeus, A., Zander, E., Marschik, P.B., Coghill, D., and Bolte, S. (2017). Annual Research Review: Quality of life and childhood mental and behavioural disorders - a critical review of the research. *J Child Psychol Psychiatry* 58, 439-469.
25. van Heijst, B.F., and Geurts, H.M. (2015). Quality of life in autism across the lifespan: a meta-analysis. *Autism* 19, 158-167.
26. Krabbenborg, L., Vissers, L.E., Schieving, J., Kleefstra, T., Kamsteeg, E.J., Veltman, J.A., Willemsen, M.A., and Van der Burg, S. (2016). Understanding the Psychosocial Effects of WES Test Results on Parents of Children with Rare Diseases. *J Genet Couns* 25, 1207-1214.
27. Vissers, L.E., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., van Lier, B., Arts, P., Wieskamp, N., del Rosario, M., et al. (2010). A de novo paradigm for mental retardation. *Nat Genet* 42, 1109-1112.
28. de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *The New England journal of medicine* 367, 1921-1929.
29. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Ende, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380, 1674-1682.
30. Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., et al. (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron* 74, 285-299.
31. Iossifov, I., O'Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nat New Biol* 515, 216-221.
32. Deciphering Developmental Disorders, S. (2015). Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 519, 223-228.
33. Deciphering Developmental Disorders, S. (2017). Prevalence and architecture of de novo mutations in developmental disorders. *Nat New Biol* 542, 433-438.
34. Willsey, A.J., Fernandez, T.V., Yu, D., King, R.A., Dietrich, A., Xing, J., Sanders, S.J., Mandell, J.D., Huang, A.Y.,

- Richer, P., et al. (2017). De Novo Coding Variants Are Strongly Associated with Tourette Disorder. *Neuron* 94, 486-499 e489.
35. Lima Lde, A., Feio-dos-Santos, A.C., Belangero, S.I., Gadelha, A., Bressan, R.A., Salum, G.A., Pan, P.M., Moriyama, T.S., Graeff-Martins, A.S., Tamanaha, A.C., et al. (2016). An integrative approach to investigate the respective roles of single-nucleotide variants and copy-number variants in Attention-Deficit/Hyperactivity Disorder. *Sci Rep* 6, 22851.
 36. Kim, D.S., Burt, A.A., Ranchalis, J.E., Wilmot, B., Smith, J.D., Patterson, K.E., Coe, B.P., Li, Y.K., Bamshad, M.J., Nikolas, M., et al. (2017). Sequencing of sporadic Attention-Deficit Hyperactivity Disorder (ADHD) identifies novel and potentially pathogenic de novo variants and excludes overlap with genes associated with autism spectrum disorder. *Am J Med Genet B Neuropsychiatr Genet* 174, 381-389.
 37. Chen, X.S., Reader, R.H., Hoischen, A., Veltman, J.A., Simpson, N.H., Francks, C., Newbury, D.F., and Fisher, S.E. (2017). Next-generation DNA sequencing identifies novel gene variants and pathways involved in specific language impairment. *Sci Rep* 7, 46105.
 38. Villanueva, P., Nudel, R., Hoischen, A., Fernandez, M.A., Simpson, N.H., Gilissen, C., Reader, R.H., Jara, L., Echeverry, M.M., Francks, C., et al. (2015). Exome sequencing in an admixed isolated population indicates NFXL1 variants confer a risk for specific language impairment. *PLoS Genet* 11, e1004925.
 39. Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nat New Biol* 506, 179-184.
 40. Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Consortium on the Genetics of, S., Group, P.S., et al. (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* 154, 518-529.
 41. Schreiber, M., Dorschner, M., and Tsuang, D. (2013). Next-generation sequencing in schizophrenia and other neuropsychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* 162B, 671-678.
 42. Hamdan, F.F., Myers, C.T., Cossette, P., Lemay, P., Spiegelman, D., Laporte, A.D., Nassif, C., Diallo, O., Monlong, J., Cadieux-Dion, M., et al. (2017). High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. *Am J Hum Genet* 101, 664-685.
 43. Epi, K.C., Epilepsy Phenome/Genome, P., Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., et al. (2013). De novo mutations in epileptic encephalopathies. *Nat New Biol* 501, 217-221.
 44. Euro, E.-R.E.S.C., Epilepsy Phenome/Genome, P., and Epi, K.C. (2014). De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. *Am J Hum Genet* 95, 360-370.
 45. Krumm, N., O'Roak, B.J., Shendure, J., and Eichler, E.E. (2014). A de novo convergence of autism genetics and molecular neuroscience. *Trends Neurosci* 37, 95-105.
 46. Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nat New Biol* 511, 344-347.
 47. Smedley, D., Schubach, M., Jacobsen, J.O., Kohler, S., Zemojtel, T., Spielmann, M., Jager, M., Hochheiser, H., Washington, N.L., McMurtry, J.A., et al. (2016). A Whole-Genome Analysis Framework for Effective Identification of Pathogenic Regulatory Variants in Mendelian Disease. *Am J Hum Genet* 99, 595-606.

48. Lionel, A.C., Costain, G., Monfared, N., Walker, S., Reuter, M.S., Hosseini, S.M., Thiruvahindrapuram, B., Merico, D., Jobling, R., Nalpathamkalam, T., et al. (2017). Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med*.
49. Monroe, G.R., Frederix, G.W., Savelberg, S.M., de Vries, T.I., Duran, K.J., van der Smagt, J.J., Terhal, P.A., van Hasselt, P.M., Kroes, H.Y., Verhoeven-Duif, N.M., et al. (2016). Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med* 18, 949-956.
50. Weinstein, V., Tanpaiboon, P., Chapman, K.A., Ah Mew, N., and Hofherr, S. (2017). Do the data really support ordering fragile X testing as a first-tier test without clinical features? *Genet Med*.
51. Hartley, T., Potter, R., Badalato, L., Smith, A.C., Jarinova, O., and Boycott, K.M. (2017). Fragile X testing as a second-tier test. *Genet Med*.
52. Maruoka, R., Takenouchi, T., Torii, C., Shimizu, A., Misu, K., Higasa, K., Matsuda, F., Ota, A., Tanito, K., Kuramochi, A., et al. (2014). The use of next-generation sequencing in molecular diagnosis of neurofibromatosis type 1: a validation study. *Genet Test Mol Biomarkers* 18, 722-735.
53. Vissers, L., van Nimwegen, K.J.M., Schieving, J.H., Kamsteeg, E.J., Kleefstra, T., Yntema, H.G., Pfundt, R., van der Wilt, G.J., Krabbenborg, L., Brunner, H.G., et al. (2017). A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genet Med* 19, 1055-1063.
54. Tan, T.Y., Dillon, O.J., Stark, Z., Schofield, D., Alam, K., Shrestha, R., Chong, B., Phelan, D., Brett, G.R., Creed, E., et al. (2017). Diagnostic Impact and Cost-effectiveness of Whole-Exome Sequencing for Ambulant Children With Suspected Monogenic Conditions. *JAMA Pediatr* 171, 855-862.
55. Valencia, C.A., Husami, A., Holle, J., Johnson, J.A., Qian, Y., Mathur, A., Wei, C., Indugula, S.R., Zou, F., Meng, H., et al. (2015). Clinical Impact and Cost-Effectiveness of Whole Exome Sequencing as a Diagnostic Tool: A Pediatric Center's Experience. *Front Pediatr* 3, 67.
56. Stark, Z., Tan, T.Y., Chong, B., Brett, G.R., Yap, P., Walsh, M., Yeung, A., Peters, H., Mordaunt, D., Cowie, S., et al. (2016). A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet Med* 18, 1090-1096.
57. Soden, S.E., Saunders, C.J., Willig, L.K., Farrow, E.G., Smith, L.D., Petrikin, J.E., LePichon, J.B., Miller, N.A., Thiffault, I., Dinwiddie, D.L., et al. (2014). Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Sci Transl Med* 6, 265ra168.
58. Green, R.C., Berg, J.S., Grody, W.W., Kalia, S.S., Korf, B.R., Martin, C.L., McGuire, A.L., Nussbaum, R.L., O'Daniel, J.M., Ormond, K.E., et al. (2013). ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 15, 565-574.
59. Amendola, L.M., Dorschner, M.O., Robertson, P.D., Salama, J.S., Hart, R., Shirts, B.H., Murray, M.L., Tokita, M.J., Gallego, C.J., Kim, D.S., et al. (2015). Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res* 25, 305-315.
60. Biesecker, L.G. (2016). Overcalling secondary findings. *Genet Med* 18, 416.
61. Dorschner, M.O., Amendola, L.M., Turner, E.H., Robertson, P.D., Shirts, B.H., Gallego, C.J., Bennett, R.L., Jones, K.L., Tokita, M.J., Bennett, J.T., et al. (2013). Actionable, pathogenic incidental findings in 1,000 participants' exomes. *Am J Hum Genet* 93, 631-640.
62. Firth, H.V., Hurst, J.A., and Hall, J.G. (2009). *Clinical genetics*. (New York: Oxford University Press).
63. Ratan, A., Miller, W., Guillory, J., Stinson, J., Seshagiri, S., and Schuster, S.C. (2013). Comparison of sequencing

- platforms for single nucleotide variant calls in a human sample. *PLoS One* 8, e55089.
64. Metzker, M.L. (2010). Sequencing technologies - the next generation. *Nat Rev Genet* 11, 31-46.
 65. Wright, C.F., FitzPatrick, D.R., and Firth, H.V. (2018). Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet*.
 66. Li, H., Ruan, J., and Durbin, R. (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18, 1851-1858.
 67. Strom, S.P., Lee, H., Das, K., Vilain, E., Nelson, S.F., Grody, W.W., and Deignan, J.L. (2014). Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. *Genet Med* 16, 510-515.
 68. Baudhuin, L.M., Lagerstedt, S.A., Klee, E.W., Fadra, N., Oglesbee, D., and Ferber, M.J. (2015). Confirming Variants in Next-Generation Sequencing Panel Testing by Sanger Sequencing. *J Mol Diagn* 17, 456-461.
 69. Genomes Project, C., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., et al. (2015). A global reference for human genetic variation. *Nat New Biol* 526, 68-74.
 70. Acuna-Hidalgo, R., Veltman, J.A., and Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biol* 17, 241.
 71. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nat New Biol* 536, 285-291.
 72. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nat New Biol* 508, 469-476.
 73. Sunyaev, S.R. (2012). Inferring causality and functional significance of human coding DNA variants. *Hum Mol Genet* 21, R10-17.
 74. Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* (80-) 347, 1260419.
 75. Thul, P.J., Akesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., Alm, T., Asplund, A., Bjork, L., Breckels, L.M., et al. (2017). A subcellular map of the human proteome. *Science* (80-) 356.
 76. Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., Keerthikumar, S., Oortveld, M.A., Kleefstra, T., Kramer, J.M., et al. (2016). Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet* 98, 149-164.
 77. Eppig, J.T., Richardson, J.E., Kadin, J.A., Ringwald, M., Blake, J.A., and Bult, C.J. (2015). Mouse Genome Informatics (MGI): reflecting on 25 years. *Mamm Genome* 26, 272-284.
 78. Brunner, H.G., and van Driel, M.A. (2004). From syndrome families to functional genomics. *Nat Rev Genet* 5, 545-551.
 79. Ng, P.C., and Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Res* 11, 863-874.
 80. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods* 7, 248-249.
 81. Schwarz, J.M., Rodelsperger, C., Schuelke, M., and Seelow, D. (2010). MutationTaster evaluates disease-causing

- potential of sequence alterations. *Nat Methods* 7, 575-576.
82. Kircher, M., Witten, D.M., Jain, P., O’Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet* 46, 310-315.
 83. Hicks, S., Wheeler, D.A., Plon, S.E., and Kimmel, M. (2011). Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. *Hum Mutat* 32, 661-668.
 84. Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., and Beroud, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37, e67.
 85. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17, 405-424.
 86. Yvonne Wallis, S.P., Ciaron McNulty, Danielle Bodmer, Erik Sistermans, Kathryn Robertson, David Moore, Stephen Abbs, Zandra Deans, Andrew Devereau. (2013). Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics. Available from: http://www.wacg-sukcom/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdfpdf.
 87. Buske, O.J., Girdea, M., Dumitriu, S., Gallinger, B., Hartley, T., Trang, H., Misyura, A., Friedman, T., Beaulieu, C., Bone, W.P., et al. (2015). PhenomeCentral: a portal for phenotypic and genotypic matchmaking of patients with rare genetic diseases. *Hum Mutat* 36, 931-940.
 88. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat* 36, 928-930.
 89. Rahbari, R., Wuster, A., Lindsay, S.J., Hardwick, R.J., Alexandrov, L.B., Turki, S.A., Dominiczak, A., Morris, A., Porteous, D., Smith, B., et al. (2016). Timing, rates and spectra of human germline mutation. *Nat Genet* 48, 126-133.
 90. Campbell, I.M., Yuan, B., Robberecht, C., Pfundt, R., Szafranski, P., McEntagart, M.E., Nagamani, S.C., Erez, A., Bartnik, M., Wisniewicka-Kowalik, B., et al. (2014). Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am J Hum Genet* 95, 173-182.
 91. Bakker, E., Van Broeckhoven, C., Bonten, E.J., van de Vooren, M.J., Veenema, H., Van Hul, W., Van Ommen, G.J., Vandenbergh, A., and Pearson, P.L. (1987). Germline mosaicism and Duchenne muscular dystrophy mutations. *Nat New Biol* 329, 554-556.
 92. Acuna-Hidalgo, R., Bo, T., Kwint, M.P., van de Vorst, M., Pinelli, M., Veltman, J.A., Hoischen, A., Vissers, L.E., and Gilissen, C. (2015). Post-zygotic Point Mutations Are an Underrecognized Source of De Novo Genomic Variation. *Am J Hum Genet* 97, 67-74.
 93. Lim, E.T., Uddin, M., De Rubeis, S., Chan, Y., Kamumbu, A.S., Zhang, X., D’Gama, A.M., Kim, S.N., Hill, R.S., Goldberg, A.P., et al. (2017). Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat Neurosci* 20, 1217-1224.
 94. Yang, Y., Muzny, D.M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 312, 1870-1879.
 95. Zhu, X., Petrovski, S., Xie, P., Ruzzo, E.K., Lu, Y.F., McSweeney, K.M., Ben-Zeev, B., Nissenkorn, A., Anikster, Y., Oz-Levi, D., et al. (2015). Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med* 17, 774-781.

96. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 369, 1502-1511.
97. Martinez, F., Caro-Llopis, A., Rosello, M., Oltra, S., Mayo, S., Monfort, S., and Orellana, C. (2017). High diagnostic yield of syndromic intellectual disability by targeted next-generation sequencing. *J Med Genet* 54, 87-92.
98. Redin, C., Gerard, B., Lauer, J., Herenger, Y., Muller, J., Quartier, A., Masurel-Paulet, A., Willems, M., Lesca, G., El-Chehadeh, S., et al. (2014). Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J Med Genet* 51, 724-736.
99. Wright, C.F., Fitzgerald, T.W., Jones, W.D., Clayton, S., McRae, J.F., van Kogelenberg, M., King, D.A., Ambridge, K., Barrett, D.M., Bayzatinova, T., et al. (2014). Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 385, 1305-1314.
100. Wright, C.F., McRae, J.F., Clayton, S., Gallone, G., Aitken, S., FitzGerald, T.W., Jones, P., Prigmore, E., Rajan, D., Lord, J., et al. (2018). Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet Med*.
101. Vissers, L.E., Gilissen, C., and Veltman, J.A. (2016). Genetic studies in intellectual disability and related disorders. *Nat Rev Genet* 17, 9-18.
102. Pfundt, R., Del Rosario, M., Vissers, L.E., Kwint, M.P., Janssen, I.M., de Leeuw, N., Yntema, H.G., Nelen, M.R., Lugtenberg, D., Kamsteeg, E.J., et al. (2016). Detection of clinically relevant copy-number variants by exome sequencing in a large cohort of genetic disorders. *Genet Med*.
103. Nambot, S., Thevenon, J., Kuentz, P., Duffourd, Y., Tisserant, E., Bruel, A.L., Mosca-Boidron, A.L., Masurel-Paulet, A., Lehalle, D., Jean-Marcais, N., et al. (2017). Clinical whole-exome sequencing for the diagnosis of rare disorders with congenital anomalies and/or intellectual disability: substantial interest of prospective annual reanalysis. *Genet Med*.
104. Need, A.C., Shashi, V., Schoch, K., Petrovski, S., and Goldstein, D.B. (2017). The importance of dynamic re-analysis in diagnostic whole exome sequencing. *J Med Genet* 54, 155-156.
105. Bergant, G., Maver, A., Lovrecic, L., Cuturilo, G., Hodzic, A., and Peterlin, B. (2017). Comprehensive use of extended exome analysis improves diagnostic yield in rare disease: a retrospective survey in 1,059 cases. *Genet Med*.
106. Eldomery, M.K., Coban-Akdemir, Z., Harel, T., Rosenfeld, J.A., Gambin, T., Stray-Pedersen, A., Kury, S., Mercier, S., Lessel, D., Denecke, J., et al. (2017). Lessons learned from additional research analyses of unsolved clinical exome cases. *Genome Med* 9, 26.
107. Baldrige, D., Heeley, J., Vineyard, M., Manwaring, L., Toler, T.L., Fassi, E., Fiala, E., Brown, S., Goss, C.W., Willing, M., et al. (2017). The Exome Clinic and the role of medical genetics expertise in the interpretation of exome sequencing results. *Genet Med* 19, 1040-1048.
108. Wenger, A.M., Guturu, H., Bernstein, J.A., and Bejerano, G. (2017). Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med* 19, 209-214.
109. Posey, J.E., Rosenfeld, J.A., James, R.A., Bainbridge, M., Niu, Z., Wang, X., Dhar, S., Wiszniewski, W., Akdemir, Z.H., Gambin, T., et al. (2016). Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genet Med* 18, 678-685.
110. Fu, F., Li, R., Li, Y., Nie, Z.Q., Lei, T.Y., Wang, D., Yang, X., Han, J., Pan, M., Zhen, L., et al. (2017). Whole exome se-

- quencing as a diagnostic adjunct to clinical testing in a tertiary referral cohort of 3988 fetuses with structural abnormalities. *Ultrasound Obstet Gynecol*.
111. Yates, C.L., Monaghan, K.G., Copenheaver, D., Retterer, K., Scuffins, J., Kucera, C.R., Friedman, B., Richard, G., and Juusola, J. (2017). Whole-exome sequencing on deceased fetuses with ultrasound anomalies: expanding our knowledge of genetic disease during fetal development. *Genet Med* 19, 1171-1178.
 112. Vora, N.L., Powell, B., Brandt, A., Strande, N., Hardisty, E., Gilmore, K., Foreman, A.K.M., Wilhelmsen, K., Bizon, C., Reilly, J., et al. (2017). Prenatal exome sequencing in anomalous fetuses: new opportunities and challenges. *Genet Med* 19, 1207-1216.
 113. Grozeva, D., Carss, K., Spasic-Boskovic, O., Tejada, M.I., Gecz, J., Shaw, M., Corbett, M., Haan, E., Thompson, E., Friend, K., et al. (2015). Targeted Next-Generation Sequencing Analysis of 1,000 Individuals with Intellectual Disability. *Hum Mutat* 36, 1197-1204.
 114. O'Roak, B.J., Stessman, H.A., Boyle, E.A., Witherspoon, K.T., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J.B., Nickerson, D.A., et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat Commun* 5, 5595.
 115. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnstrom, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. *Nat Genet* 46, 944-950.
 116. Luscan, A., Laurendeau, I., Malan, V., Francannet, C., Odent, S., Giuliano, F., Lacombe, D., Touraine, R., Vidaud, M., Pasmant, E., et al. (2014). Mutations in SETD2 cause a novel overgrowth condition. *J Med Genet* 51, 512-517.
 117. Lumish, H.S., Wynn, J., Devinsky, O., and Chung, W.K. (2015). Brief Report: SETD2 Mutation in a Child with Autism, Intellectual Disabilities and Epilepsy. *J Autism Dev Disord* 45, 3764-3770.
 118. Neveling, K., Feenstra, I., Gilissen, C., Hoefsloot, L.H., Kamsteeg, E.J., Mensenkamp, A.R., Rodenburg, R.J., Yntema, H.G., Spruijt, L., Vermeer, S., et al. (2013). A post-hoc comparison of the utility of sanger sequencing and exome sequencing for the diagnosis of heterogeneous diseases. *Hum Mutat* 34, 1721-1726.
 119. Genome Diagnostics Nijmegen. Gene Panel: Intellectual Disability <https://www.radboudumc.nl/Informatievoorverwijzers/Genoomdiagnostiek/en/Pages/Intellectualdisability.aspx> (2015).
 120. Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., et al. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nat New Biol* 488, 471-475.
 121. Goeman, J.J., and Solari, A. (2014). Multiple hypothesis testing in genomics. *Stat Med* 33, 1946-1978.
 122. MacArthur, D.G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., Jostins, L., Habegger, L., Pickrell, J.K., Montgomery, S.B., et al. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science* (80-) 335, 823-828.
 123. Zhu, J., He, F., Song, S., Wang, J., and Yu, J. (2008). How many human genes can be defined as housekeeping with current expression data? *BMC Genomics* 9, 172.
 124. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet* 9, e1003709.
 125. Mefford, H.C., Batshaw, M.L., and Hoffman, E.P. (2012). Genomics, intellectual disability, and autism. *N Engl J Med* 366, 733-743.

126. Lelieveld, S.H., Reijnders, M.R., Pfundt, R., Yntema, H.G., Kamsteeg, E.J., de Vries, P., de Vries, B.B., Willemsen, M.H., Kleefstra, T., Lohner, K., et al. (2016). Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci* 19, 1194-1196.
127. Sharma, A., Takata, H., Shibahara, K., Bubulya, A., and Bubulya, P.A. (2010). Son is essential for nuclear speckle organization and cell cycle progression. *Mol Biol Cell* 21, 650-663.
128. Ahn, E.Y., DeKever, R.C., Lo, M.C., Nguyen, T.A., Matsuura, S., Boyapati, A., Pandit, S., Fu, X.D., and Zhang, D.E. (2011). SON controls cell-cycle progression by coordinated regulation of RNA splicing. *Mol Cell* 42, 185-198.
129. Hickey, C.J., Kim, J.H., and Ahn, E.Y. (2014). New discoveries of old SON: a link between RNA splicing and cancer. *J Cell Biochem* 115, 224-231.
130. Genome of the Netherlands, C. (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet* 46, 818-825.
131. Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* 44, 1365-1369.
132. Petrovski, S., Gussow, A.B., Wang, Q., Halvorsen, M., Han, Y., Weir, W.H., Allen, A.S., and Goldstein, D.B. (2015). The Intolerance of Regulatory Sequence to Genetic Variation Predicts Gene Dosage Sensitivity. *PLoS Genet* 11, e1005492.
133. Sharma, A., Markey, M., Torres-Munoz, K., Varia, S., Kadakia, M., Bubulya, A., and Bubulya, P.A. (2011). Son maintains accurate splicing for a subset of human pre-mRNAs. *J Cell Sci* 124, 4286-4298.
134. Kim, J.H., Baddoo, M.C., Park, E.Y., Stone, J.K., Park, H., Butler, T.W., Huang, G., Yan, X., Pauli-Behn, F., Myers, R.M., et al. (2016). SON and Its Alternatively Spliced Isoforms Control MLL Complex-Mediated H3K4me3 and Transcription of Leukemia-Associated Genes. *Mol Cell* 61, 859-873.
135. Ahn, E.E., Higashi, T., Yan, M., Matsuura, S., Hickey, C.J., Lo, M.C., Shia, W.J., DeKever, R.C., and Zhang, D.E. (2013). SON protein regulates GATA-2 through transcriptional control of the microRNA 23a~27a~24~2 cluster. *J Biol Chem* 288, 5381-5388.
136. Lu, X., Goke, J., Sachs, F., Jacques, P.E., Liang, H., Feng, B., Bourque, G., Bubulya, P.A., and Ng, H.H. (2013). SON connects the splicing-regulatory network with pluripotency in human embryonic stem cells. *Nat Cell Biol* 15, 1141-1152.
137. Bilguvar, K., Ozturk, A.K., Louvi, A., Kwan, K.Y., Choi, M., Tatli, B., Yalnizoglu, D., Tuysuz, B., Caglayan, A.O., Gokben, S., et al. (2010). Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nat New Biol* 467, 207-210.
138. Chen, J.F., Zhang, Y., Wilde, J., Hansen, K.C., Lai, F., and Niswander, L. (2014). Microcephaly disease gene Wdr62 regulates mitotic progression of embryonic neural stem cells and brain size. *Nat Commun* 5, 3885.
139. Jamuar, S.S., Lam, A.T., Kircher, M., D'Gama, A.M., Wang, J., Barry, B.J., Zhang, X., Hill, R.S., Partlow, J.N., Rozzo, A., et al. (2014). Somatic mutations in cerebral cortical malformations. *N Engl J Med* 371, 733-743.
140. Jensen, L.R., Lenzner, S., Moser, B., Freude, K., Tzschach, A., Wei, C., Fryns, J.P., Chelly, J., Turner, G., Moraine, C., et al. (2007). X-linked mental retardation: a comprehensive molecular screen of 47 candidate genes from a 7.4 Mb interval in Xp11. *Eur J Hum Genet* 15, 68-75.
141. Nicholas, A.K., Khurshid, M., Desir, J., Carvalho, O.P., Cox, J.J., Thornton, G., Kausar, R., Ansar, M., Ahmad, W., Verloes, A., et al. (2010). WDR62 is associated with the spindle pole and is mutated in human microcephaly.

- Nat Genet 42, 1010-1014.
142. Nofech-Mozes, Y., Blaser, S.I., Kobayashi, J., Grunebaum, E., and Roifman, C.M. (2007). Neurologic abnormalities in patients with adenosine deaminase deficiency. *Pediatr Neurol* 37, 218-221.
 143. Poirier, K., Lebrun, N., Broix, L., Tian, G., Saillour, Y., Boscheron, C., Parrini, E., Valence, S., Pierre, B.S., Oger, M., et al. (2013). Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. *Nat Genet* 45, 639-647.
 144. Shen, J., Gilmore, E.C., Marshall, C.A., Haddadin, M., Reynolds, J.J., Eyaid, W., Bodell, A., Barry, B., Gleason, D., Allen, K., et al. (2010). Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. *Nat Genet* 42, 245-249.
 145. Yu, T.W., Mochida, G.H., Tischfield, D.J., Sgaier, S.K., Flores-Sarnat, L., Sergi, C.M., Topcu, M., McDonald, M.T., Barry, B.J., Felie, J.M., et al. (2010). Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. *Nat Genet* 42, 1015-1020.
 146. Fox, J.W., Lamperti, E.D., Eksioğlu, Y.Z., Hong, S.E., Feng, Y., Graham, D.A., Scheffer, I.E., Dobyns, W.B., Hirsch, B.A., Radtke, R.A., et al. (1998). Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotopia. *Neuron* 21, 1315-1325.
 147. Sillje, H.H., Takahashi, K., Tanaka, K., Van Houwe, G., and Nigg, E.A. (1999). Mammalian homologues of the plant Tousled gene code for cell-cycle-regulated kinases with maximal activities linked to ongoing DNA replication. *EMBO J* 18, 5691-5702.
 148. Yamakawa, A., Kameoka, Y., Hashimoto, K., Yoshitake, Y., Nishikawa, K., Tanihara, K., and Date, T. (1997). cDNA cloning and chromosomal mapping of genes encoding novel protein kinases termed PKU-alpha and PKU-beta, which have nuclear localization signal. *Gene* 202, 193-201.
 149. Groth, A., Lukas, J., Nigg, E.A., Sillje, H.H., Wernstedt, C., Bartek, J., and Hansen, K. (2003). Human Tousled like kinases are targeted by an ATM- and Chk1-dependent DNA damage checkpoint. *EMBO J* 22, 1676-1687.
 150. Sillje, H.H., and Nigg, E.A. (2001). Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases. *Curr Biol* 11, 1068-1073.
 151. Carrera, P., Moshkin, Y.M., Gronke, S., Sillje, H.H., Nigg, E.A., Jackle, H., and Karch, F. (2003). Tousled-like kinase functions with the chromatin assembly pathway regulating nuclear divisions. *Genes Dev* 17, 2578-2590.
 152. Li, Y., DeFatta, R., Anthony, C., Sunavala, G., and De Benedetti, A. (2001). A translationally regulated Tousled kinase phosphorylates histone H3 and confers radioresistance when overexpressed. *Oncogene* 20, 726-738.
 153. Klimovskaia, I.M., Young, C., Stromme, C.B., Menard, P., Jasencakova, Z., Mejlvang, J., Ask, K., Ploug, M., Nielsen, M.L., Jensen, O.N., et al. (2014). Tousled-like kinases phosphorylate Asf1 to promote histone supply during DNA replication. *Nat Commun* 5, 3394.
 154. Bruinsma, W., van den Berg, J., Aprelia, M., and Medema, R.H. (2016). Tousled-like kinase 2 regulates recovery from a DNA damage-induced G2 arrest. *EMBO Rep* 17, 659-670.
 155. Segura-Bayona, S., Knobel, P.A., Gonzalez-Buron, H., Youssef, S.A., Pena-Blanco, A., Coyaude, E., Lopez-Rovira, T., Rein, K., Palenzuela, L., Colombelli, J., et al. (2017). Differential requirements for Tousled-like kinases 1 and 2 in mammalian development. *Cell Death Differ*.
 156. Ishigaki, Y., Li, X., Serin, G., and Maquat, L.E. (2001). Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* 106, 607-617.

157. Hashimoto, M., Matsui, T., Iwabuchi, K., and Date, T. (2008). PKU-beta/TLK1 regulates myosin II activities, and is required for accurate equaled chromosome segregation. *Mutat Res* 657, 63-67.
158. De Rubéis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nat New Biol* 515, 209-215.
159. Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* (80-) 350, 1262-1266.
160. Rauen, K.A. (2013). The RASopathies. *Annual review of genomics and human genetics* 14, 355-369.
161. Van Aelst, L., and D'Souza-Schorey, C. (1997). Rho GTPases and signaling networks. *Genes & development* 11, 2295-2322.
162. Heasman, S.J., and Ridley, A.J. (2008). Mammalian Rho GTPases: new insights into their functions from in vivo studies. *Nature reviews Molecular cell biology* 9, 690-701.
163. Duquette, P.M., and Lamarche-Vane, N. (2014). Rho GTPases in embryonic development. *Small GTPases* 5, 8.
164. Stankiewicz, T.R., and Linseman, D.A. (2014). Rho family GTPases: key players in neuronal development, neuronal survival, and neurodegeneration. *Frontiers in cellular neuroscience* 8, 314.
165. Aspenstrom, P., Fransson, A., and Saras, J. (2004). Rho GTPases have diverse effects on the organization of the actin filament system. *The Biochemical journal* 377, 327-337.
166. Boureux, A., Vignal, E., Faure, S., and Fort, P. (2007). Evolution of the Rho family of ras-like GTPases in eukaryotes. *Molecular biology and evolution* 24, 203-216.
167. Tahirovic, S., Hellal, F., Neukirchen, D., Hindges, R., Garvalov, B.K., Flynn, K.C., Stradal, T.E., Chrostek-Grashoff, A., Brakebusch, C., and Bradke, F. (2010). Rac1 regulates neuronal polarization through the WAVE complex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 30, 6930-6943.
168. Wojnacki, J., Quassollo, G., Marzolo, M.P., and Caceres, A. (2014). Rho GTPases at the crossroad of signaling networks in mammals: impact of Rho-GTPases on microtubule organization and dynamics. *Small GTPases* 5, e28430.
169. Saci, A., Cantley, L.C., and Carpenter, C.L. (2011). Rac1 regulates the activity of mTORC1 and mTORC2 and controls cellular size. *Molecular cell* 42, 50-61.
170. Sugihara, K., Nakatsuji, N., Nakamura, K., Nakao, K., Hashimoto, R., Otani, H., Sakagami, H., Kondo, H., Nozawa, S., Aiba, A., et al. (1998). Rac1 is required for the formation of three germ layers during gastrulation. *Oncogene* 17, 3427-3433.
171. Threadgill, R., Bobb, K., and Ghosh, A. (1997). Regulation of dendritic growth and remodeling by Rho, Rac, and Cdc42. *Neuron* 19, 625-634.
172. Leone, D.P., Srinivasan, K., Brakebusch, C., and McConnell, S.K. (2010). The rho GTPase Rac1 is required for proliferation and survival of progenitors in the developing forebrain. *Developmental neurobiology* 70, 659-678.
173. Chen, L., Melendez, J., Campbell, K., Kuan, C.Y., and Zheng, Y. (2009). Rac1 deficiency in the forebrain results in neural progenitor reduction and microcephaly. *Developmental biology* 325, 162-170.
174. Matos, P., Skaug, J., Marques, B., Beck, S., Verissimo, F., Gespach, C., Boavida, M.G., Scherer, S.W., and Jordan, P. (2000). Small GTPase Rac1: structure, localization, and expression of the human gene. *Biochemical and biophysical research communications* 277, 741-751.

175. Jordan, P., Brazao, R., Boavida, M.G., Gespach, C., and Chastre, E. (1999). Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors. *Oncogene* 18, 6835-6839.
176. Colicelli, J. (2004). Human RAS superfamily proteins and related GTPases. *Science's STKE : signal transduction knowledge environment* 2004, RE13.
177. Hobbs, G.A., Mitchell, L.E., Arrington, M.E., Gunawardena, H.P., DeCristo, M.J., Loeser, R.F., Chen, X., Cox, A.D., and Campbell, S.L. (2015). Redox regulation of Rac1 by thiol oxidation. *Free radical biology & medicine* 79, 237-250.
178. Gao, Y., Xing, J., Streuli, M., Leto, T.L., and Zheng, Y. (2001). Trp(56) of rac1 specifies interaction with a subset of guanine nucleotide exchange factors. *The Journal of biological chemistry* 276, 47530-47541.
179. De Baets, G., Van Durme, J., Reumers, J., Maurer-Stroh, S., Vanhee, P., Dopazo, J., Schymkowitz, J., and Rousseau, F. (2012). SNPEffect 4.0: on-line prediction of molecular and structural effects of protein-coding variants. *Nucleic acids research* 40, D935-939.
180. Price, L.S., Leng, J., Schwartz, M.A., and Bokoch, G.M. (1998). Activation of Rac and Cdc42 by integrins mediates cell spreading. *Molecular biology of the cell* 9, 1863-1871.
181. Ridley, A.J., Paterson, H.F., Johnston, C.L., Diekmann, D., and Hall, A. (1992). The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* 70, 401-410.
182. Machesky, L.M., and Hall, A. (1997). Role of actin polymerization and adhesion to extracellular matrix in Rac- and Rho-induced cytoskeletal reorganization. *The Journal of cell biology* 138, 913-926.
183. Michiels, F., Habets, G.G., Stam, J.C., van der Kammen, R.A., and Collard, J.G. (1995). A role for Rac in Tiam1-induced membrane ruffling and invasion. *Nature* 375, 338-340.
184. Jao, L.E., Wentz, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci U S A* 110, 13904-13909.
185. Perles, Z., Moon, S., Ta-Shma, A., Yaacov, B., Francescato, L., Edvardson, S., Rein, A.J., Elpeleg, O., and Katsanis, N. (2015). A human laterality disorder caused by a homozygous deleterious mutation in MMP21. *J Med Genet* 52, 840-847.
186. Mulherkar, S., Uddin, M.D., Couvillon, A.D., Sillitoe, R.V., and Tolia, K.F. (2014). The small GTPases RhoA and Rac1 regulate cerebellar development by controlling cell morphogenesis, migration and foliation. *Developmental biology* 394, 39-53.
187. Nevado, J., Rosenfeld, J.A., Mena, R., Palomares-Bravo, M., Vallespin, E., Angeles Mori, M., Tenorio, J.A., Gripp, K.W., Denenberg, E., Del Campo, M., et al. (2015). PIAS4 is associated with macro/microcephaly in the novel interstitial 19p13.3 microdeletion/microduplication syndrome. *European journal of human genetics : EJHG* 23, 1615-1626.
188. Van Dijk, A., van der Werf, I.M., Reyniers, E., Scheers, S., Azage, M., Siefkas, K., Van der Aa, N., Lacroix, A., Rosenfeld, J., Argiropoulos, B., et al. (2015). Five patients with a chromosome 1q21.1 triplication show macrocephaly, increased weight and facial similarities. *European journal of medical genetics* 58, 503-508.
189. Shinawi, M., Liu, P., Kang, S.H., Shen, J., Belmont, J.W., Scott, D.A., Probst, F.J., Craigen, W.J., Graham, B.H., Pursley, A., et al. (2010). Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *Journal of medical genetics* 47, 332-341.
190. Riviere, J.B., Mirzaa, G.M., O'Roak, B.J., Beddaoui, M., Alcantara, D., Conway, R.L., St-Onge, J., Schwartz-

- truber, J.A., Gripp, K.W., Nikkel, S.M., et al. (2012). De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat Genet* 44, 934-940.
191. Rosner, M., Hanneder, M., Siegel, N., Valli, A., Fuchs, C., and Hengstschlager, M. (2008). The mTOR pathway and its role in human genetic diseases. *Mutation research* 659, 284-292.
 192. Hetmanski, J.H., Schwartz, J.M., and Caswell, P.T. (2016). Rationalizing Rac1 and RhoA GTPase signaling: A mathematical approach. *Small GTPases*, 1-6.
 193. Pengelly, R.J., Greville-Heygate, S., Schmidt, S., Seaby, E.G., Jabalameli, M.R., Mehta, S.G., Parker, M.J., Goudie, D., Fagotto-Kaufmann, C., Mercer, C., et al. (2016). Mutations specific to the Rac-GEF domain of TRIO cause intellectual disability and microcephaly. *J Med Genet*.
 194. Ba, W., Yan, Y., Reijnders, M.R., Schuurs-Hoeijmakers, J.H., Feenstra, I., Bongers, E.M., Bosch, D.G., De Leeuw, N., Pfundt, R., Gilissen, C., et al. (2016). TRIO loss of function is associated with mild intellectual disability and affects dendritic branching and synapse function. *Hum Mol Genet* 25, 892-902.
 195. Hollstein, R., Parry, D.A., Nalbach, L., Logan, C.V., Strom, T.M., Hartill, V.L., Carr, I.M., Korenke, G.C., Uppal, S., Ahmed, M., et al. (2015). HACE1 deficiency causes an autosomal recessive neurodevelopmental syndrome. *Journal of medical genetics* 52, 797-803.
 196. Peper, J.S., Brouwer, R.M., Boomsma, D.I., Kahn, R.S., and Hulshoff Pol, H.E. (2007). Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum Brain Mapp* 28, 464-473.
 197. Hibar, D.P., Stein, J.L., Renteria, M.E., Arias-Vasquez, A., Desrivieres, S., Jahanshad, N., Toro, R., Wittfeld, K., Abramovic, L., Andersson, M., et al. (2015). Common genetic variants influence human subcortical brain structures. *Nat New Biol* 520, 224-229.
 198. Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. *Cell* 169, 361-371.
 199. Tee, A.R., Manning, B.D., Roux, P.P., Cantley, L.C., and Blenis, J. (2003). Tuberous sclerosis complex gene products, Tuberin and Hamartin, control mTOR signaling by acting as a GTPase-activating protein complex toward Rheb. *Curr Biol* 13, 1259-1268.
 200. Kwon, C.H., Luikart, B.W., Powell, C.M., Zhou, J., Matheny, S.A., Zhang, W., Li, Y., Baker, S.J., and Parada, L.F. (2006). Pten regulates neuronal arborization and social interaction in mice. *Neuron* 50, 377-388.
 201. Li, Y.H., Werner, H., and Puschel, A.W. (2008). Rheb and mTOR regulate neuronal polarity through Rap1B. *J Biol Chem* 283, 33784-33792.
 202. Nie, D., Di Nardo, A., Han, J.M., Baharanyi, H., Kramvis, I., Huynh, T., Dabora, S., Codeluppi, S., Pandolfi, P.P., Pasquale, E.B., et al. (2010). Tsc2-Rheb signaling regulates EphA-mediated axon guidance. *Nat Neurosci* 13, 163-172.
 203. Urbanska, M., Gozdz, A., Swiech, L.J., and Jaworski, J. (2012). Mammalian target of rapamycin complex 1 (mTORC1) and 2 (mTORC2) control the dendritic arbor morphology of hippocampal neurons. *J Biol Chem* 287, 30240-30256.
 204. Hartman, N.W., Lin, T.V., Zhang, L., Paquette, G.E., Feliciano, D.M., and Bordey, A. (2013). mTORC1 targets the translational repressor 4E-BP2, but not S6 kinase 1/2, to regulate neural stem cell self-renewal in vivo. *Cell Rep* 5, 433-444.
 205. Tavazoie, S.F., Alvarez, V.A., Ridenour, D.A., Kwiatkowski, D.J., and Sabatini, B.L. (2005). Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nat Neurosci* 8, 1727-1734.
 206. Feliciano, D.M., Su, T., Lopez, J., Platel, J.C., and Bordey, A. (2011). Single-cell Tsc1 knockout during corticogene-

- sis generates tuber-like lesions and reduces seizure threshold in mice. *J Clin Invest* 121, 1596-1607.
207. Zhou, J., Shrikhande, G., Xu, J., McKay, R.M., Burns, D.K., Johnson, J.E., and Parada, L.F. (2011). Tsc1 mutant neural stem/progenitor cells exhibit migration deficits and give rise to subependymal lesions in the lateral ventricle. *Genes Dev* 25, 1595-1600.
 208. Baek, S.T., Copeland, B., Yun, E.J., Kwon, S.K., Guemez-Gamboa, A., Schaffer, A.E., Kim, S., Kang, H.C., Song, S., Mathern, G.W., et al. (2015). An AKT3-FOXG1-reelin network underlies defective migration in human focal malformations of cortical development. *Nat Med* 21, 1445-1454.
 209. Lin, T.V., Hsieh, L., Kimura, T., Malone, T.J., and Bordey, A. (2016). Normalizing translation through 4E-BP prevents mTOR-driven cortical mislamination and ameliorates aberrant neuron integration. *Proc Natl Acad Sci U S A* 113, 11330-11335.
 210. Moon, U.Y., Park, J.Y., Park, R., Cho, J.Y., Hughes, L.J., McKenna, J., 3rd, Goetzl, L., Cho, S.H., Crino, P.B., Gambello, M.J., et al. (2015). Impaired Reelin-Dab1 Signaling Contributes to Neuronal Migration Deficits of Tuberous Sclerosis Complex. *Cell Rep* 12, 965-978.
 211. Lafourcade, C.A., Lin, T.V., Feliciano, D.M., Zhang, L., Hsieh, L.S., and Bordey, A. (2013). Rheb activation in sub-ventricular zone progenitors leads to heterotopia, ectopic neuronal differentiation, and rapamycin-sensitive olfactory micronodules and dendrite hypertrophy of newborn neurons. *J Neurosci* 33, 2419-2431.
 212. Bateup, H.S., Takasaki, K.T., Saulnier, J.L., Deneffrio, C.L., and Sabatini, B.L. (2011). Loss of Tsc1 in vivo impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J Neurosci* 31, 8862-8869.
 213. Sugiura, H., Yasuda, S., Katsurabayashi, S., Kawano, H., Endo, K., Takasaki, K., Iwasaki, K., Ichikawa, M., Kobayashi, T., Hino, O., et al. (2015). Rheb activation disrupts spine synapse formation through accumulation of syntenin in tuberous sclerosis complex. *Nat Commun* 6, 6842.
 214. Tang, G., Gudsnek, K., Kuo, S.H., Cotrina, M.L., Rosoklija, G., Sosunov, A., Sonders, M.S., Kanter, E., Castagna, C., Yamamoto, A., et al. (2014). Loss of mTOR-dependent macroautophagy causes autistic-like synaptic pruning deficits. *Neuron* 83, 1131-1143.
 215. Lozovaya, N., Gataullina, S., Tsintsadze, T., Tsintsadze, V., Pallesi-Pocachard, E., Minlebaev, M., Goriounova, N.A., Buhler, E., Watrin, F., Shityakov, S., et al. (2014). Selective suppression of excessive GluN2C expression rescues early epilepsy in a tuberous sclerosis murine model. *Nat Commun* 5, 4563.
 216. Hoeffer, C.A., Santini, E., Ma, T., Arnold, E.C., Whelan, A.M., Wong, H., Pierre, P., Pelletier, J., and Klann, E. (2013). Multiple components of eIF4F are required for protein synthesis-dependent hippocampal long-term potentiation. *J Neurophysiol* 109, 68-76.
 217. Santini, E., Huynh, T.N., MacAskill, A.F., Carter, A.G., Pierre, P., Ruggero, D., Kaphzan, H., and Klann, E. (2013). Exaggerated translation causes synaptic and behavioural aberrations associated with autism. *Nat New Biol* 493, 411-415.
 218. Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. *Nat New Biol* 485, 109-113.
 219. Raab-Graham, K.F., Haddick, P.C., Jan, Y.N., and Jan, L.Y. (2006). Activity- and mTOR-dependent suppression of Kv1.1 channel mRNA translation in dendrites. *Science* (80-) 314, 144-148.
 220. Huang, W.C., Chen, Y., and Page, D.T. (2016). Hyperconnectivity of prefrontal cortex to amygdala projections in a mouse model of macrocephaly/autism syndrome. *Nat Commun* 7, 13421.
 221. Zeng, L.H., Rensing, N.R., Zhang, B., Gutmann, D.H., Gambello, M.J., and Wong, M. (2011). Tsc2 gene inactivation

- causes a more severe epilepsy phenotype than Tsc1 inactivation in a mouse model of tuberous sclerosis complex. *Hum Mol Genet* 20, 445-454.
222. Abs, E., Goorden, S.M., Schreiber, J., Overwater, I.E., Hoogveen-Westerveld, M., Bruinsma, C.F., Aganovic, E., Borgesius, N.Z., Nellist, M., and Elgersma, Y. (2013). TORC1-dependent epilepsy caused by acute biallelic Tsc1 deletion in adult mice. *Ann Neurol* 74, 569-579.
 223. Pun, R.Y., Rolle, I.J., Lasarge, C.L., Hosford, B.E., Rosen, J.M., Uhl, J.D., Schmeltzer, S.N., Faulkner, C., Bronson, S.L., Murphy, B.L., et al. (2012). Excessive activation of mTOR in postnatally generated granule cells is sufficient to cause epilepsy. *Neuron* 75, 1022-1034.
 224. Hsieh, L.S., Wen, J.H., Claycomb, K., Huang, Y., Harrsch, F.A., Naegele, J.R., Hyder, F., Buchanan, G.F., and Borden, A. (2016). Convulsive seizures from experimental focal cortical dysplasia occur independently of cell misplacement. *Nat Commun* 7, 11753.
 225. Goorden, S.M., van Woerden, G.M., van der Weerd, L., Cheadle, J.P., and Elgersma, Y. (2007). Cognitive deficits in Tsc1+/- mice in the absence of cerebral lesions and seizures. *Ann Neurol* 62, 648-655.
 226. Goorden, S.M., Abs, E., Bruinsma, C.F., Riemsdijk, F.W., van Woerden, G.M., and Elgersma, Y. (2015). Intact neuronal function in Rheb1 mutant mice: implications for TORC1-based treatments. *Hum Mol Genet* 24, 3390-3398.
 227. Franz, D.N., Belousova, E., Sparagana, S., Bebin, E.M., Frost, M., Kuperman, R., Witt, O., Kohrman, M.H., Flamini, J.R., Wu, J.Y., et al. (2013). Efficacy and safety of everolimus for subependymal giant cell astrocytomas associated with tuberous sclerosis complex (EXIST-1): a multicentre, randomised, placebo-controlled phase 3 trial. *Lancet* 381, 125-132.
 228. Overwater, I.E., Rietman, A.B., Bindels-de Heus, K., Looman, C.W., Rizopoulos, D., Sibindi, T.M., Cherian, P.J., Jansen, F.E., Moll, H.A., Elgersma, Y., et al. (2016). Sirolimus for epilepsy in children with tuberous sclerosis complex: A randomized controlled trial. *Neurology* 87, 1011-1018.
 229. French, J.A., Lawson, J.A., Yapici, Z., Ikeda, H., Polster, T., Nabbout, R., Curatolo, P., de Vries, P.J., Dlugos, D.J., Berkowitz, N., et al. (2016). 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* 388, 2153-2163.
 230. Wheless, J.W. (2015). Use of the mTOR inhibitor everolimus in a patient with multiple manifestations of tuberous sclerosis complex including epilepsy. *Epilepsy & behavior case reports* 4, 63-66.
 231. Talos, D.M., Sun, H., Zhou, X., Fitzgerald, E.C., Jackson, M.C., Klein, P.M., Lan, V.J., Joseph, A., and Jensen, F.E. (2012). The interaction between early life epilepsy and autistic-like behavioral consequences: a role for the mammalian target of rapamycin (mTOR) pathway. *PLoS One* 7, e35885.
 232. Lee, J.H., Huynh, M., Silhavy, J.L., Kim, S., Dixon-Salazar, T., Heiberg, A., Scott, E., Bafna, V., Hill, K.J., Collazo, A., et al. (2012). De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nat Genet* 44, 941-945.
 233. Mirzaa, G.M., Campbell, C.D., Solovieff, N., Goold, C.P., Jansen, L.A., Menon, S., Timms, A.E., Conti, V., Biag, J.D., Olds, C., et al. (2016). Association of MTOR Mutations With Developmental Brain Disorders, Including Megalencephaly, Focal Cortical Dysplasia, and Pigmentary Mosaicism. *JAMA neurology* 73, 836-845.
 234. Lim, J.S., Kim, W.I., Kang, H.C., Kim, S.H., Park, A.H., Park, E.K., Cho, Y.W., Kim, S., Kim, H.M., Kim, J.A., et al. (2015). Brain somatic mutations in MTOR cause focal cortical dysplasia type II leading to intractable epilep-

- sy. *Nat Med* 21, 395-400.
235. Adams, H.H., Hibar, D.P., Chouraki, V., Stein, J.L., Nyquist, P.A., Renteria, M.E., Trompet, S., Arias-Vasquez, A., Seshadri, S., Desrivieres, S., et al. (2016). Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat Neurosci*.
 236. Takei, N., and Nawa, H. (2014). mTOR signaling and its roles in normal and abnormal brain development. *Front Mol Neurosci* 7, 28.
 237. Kassai, H., Sugaya, Y., Noda, S., Nakao, K., Maeda, T., Kano, M., and Aiba, A. (2014). Selective activation of mTORC1 signaling recapitulates microcephaly, tuberous sclerosis, and neurodegenerative diseases. *Cell Rep* 7, 1626-1639.
 238. Dehay, C., and Kennedy, H. (2007). Cell-cycle control and cortical development. *Nat Rev Neurosci* 8, 438-450.
 239. Molyneaux, B.J., Arlotta, P., Menezes, J.R., and Macklis, J.D. (2007). Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* 8, 427-437.
 240. Guerini, R., and Parrini, E. (2010). Neuronal migration disorders. *Neurobiol Dis* 38, 154-166.
 241. Cho, C.H. (2011). Frontier of epilepsy research - mTOR signaling pathway. *Exp Mol Med* 43, 231-274.
 242. Houge, G., Haesen, D., Vissers, L.E., Mehta, S., Parker, M.J., Wright, M., Vogt, J., McKee, S., Tolmie, J.L., Cord-eiro, N., et al. (2015). B56delta-related protein phosphatase 2A dysfunction identified in patients with intellectual disability. *J Clin Invest* 125, 3051-3062.
 243. Pers, T.H. (2016). Gene set analysis for interpreting genetic studies. *Hum Mol Genet* 25, R133-R140.
 244. Sun, T., and Hevner, R.F. (2014). Growth and folding of the mammalian cerebral cortex: from molecules to malformations. *Nat Rev Neurosci* 15, 217-232.
 245. Meikle, L., Pollizzi, K., Egnor, A., Kramvis, I., Lane, H., Sahin, M., and Kwiatkowski, D.J. (2008). Response of a neuronal model of tuberous sclerosis to mammalian target of rapamycin (mTOR) inhibitors: effects on mTORC1 and Akt signaling lead to improved survival and function. *J Neurosci* 28, 5422-5432.
 246. Hanai, S., Sukigara, S., Dai, H., Owa, T., Horike, S.I., Otsuki, T., Saito, T., Nakagawa, E., Ikegaya, N., Kaido, T., et al. (2017). Pathologic Active mTOR Mutation in Brain Malformation with Intractable Epilepsy Leads to Cell-Autonomous Migration Delay. *Am J Pathol* 187, 1177-1185.
 247. Loconte, D.C., Grossi, V., Bozzao, C., Forte, G., Bagnulo, R., Stella, A., Lastella, P., Cutrone, M., Benedicenti, F., Susca, F.C., et al. (2015). Molecular and Functional Characterization of Three Different Postzygotic Mutations in PIK3CA-Related Overgrowth Spectrum (PROS) Patients: Effects on PI3K/AKT/mTOR Signaling and Sensitivity to PIK3 Inhibitors. *PLoS One* 10, e0123092.
 248. Fang, Y., Westbrook, R., Hill, C., Boparai, R.K., Arum, O., Spong, A., Wang, F., Javors, M.A., Chen, J., Sun, L.Y., et al. (2013). Duration of rapamycin treatment has differential effects on metabolism in mice. *Cell Metab* 17, 456-462.
 249. Laplante, M., and Sabatini, D.M. (2012). mTOR signaling in growth control and disease. *Cell* 149, 274-293.
 250. Shimobayashi, M., and Hall, M.N. (2014). Making new contacts: the mTOR network in metabolism and signalling crosstalk. *Nat Rev Mol Cell Biol* 15, 155-162.
 251. Kolch, W. (2005). Coordinating ERK/MAPK signalling through scaffolds and inhibitors. *Nat Rev Mol Cell Biol* 6, 827-837.
 252. Turner, T.N., Yi, Q., Krumm, N., Huddleston, J., Hoekzema, K., HA, F.S., Doebley, A.L., Bernier, R.A., Nickerson, D.A., and Eichler, E.E. (2017). denovo-db: a compendium of human de novo variants. *Nucleic Acids Res* 45,

D804-D811.

253. Thompson, P.M., Stein, J.L., Medland, S.E., Hibar, D.P., Vasquez, A.A., Renteria, M.E., Toro, R., Jahanshad, N., Schumann, G., Franke, B., et al. (2014). The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav* 8, 153-182.
254. Psaty, B.M., O'Donnell, C.J., Gudnason, V., Lunetta, K.L., Folsom, A.R., Rotter, J.I., Uitterlinden, A.G., Harris, T.B., Witteman, J.C., Boerwinkle, E., et al. (2009). Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. *Circ Cardiovasc Genet* 2, 73-80.
255. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 11, e1004219.
256. Genomes Project, C., Abecasis, G.R., Altshuler, D., Auton, A., Brooks, L.D., Durbin, R.M., Gibbs, R.A., Hurles, M.E., and McVean, G.A. (2010). A map of human genome variation from population-scale sequencing. *Nat New Biol* 467, 1061-1073.
257. de Leeuw, C.A., Neale, B.M., Heskes, T., and Posthuma, D. (2016). The statistical properties of gene-set analysis. *Nat Rev Genet* 17, 353-364.
258. Cawthon, R.M., O'Connell, P., Buchberg, A.M., Viskochil, D., Weiss, R.B., Culver, M., Stevens, J., Jenkins, N.A., Copeland, N.G., and White, R. (1990). Identification and characterization of transcripts from the neurofibromatosis 1 region: the sequence and genomic structure of EVI2 and mapping of other transcripts. *Genomics* 7, 555-565.
259. Goslink, G., Banker, G. (1991). *Culturing Nerve cells.*(MIT Press).
260. Saito, T., and Nakatsuji, N. (2001). Efficient gene transfer into the embryonic mouse brain using in vivo electroporation. *Dev Biol* 240, 237-246.
261. Taniguchi, Y., Young-Pearse, T., Sawa, A., and Kamiya, A. (2012). In utero electroporation as a tool for genetic manipulation in vivo to study psychiatric disorders: from genes to circuits and behaviors. *Neuroscientist* 18, 169-179.
262. Wandinger-Ness, A., and Zerial, M. (2014). Rab proteins and the compartmentalization of the endosomal system. *Cold Spring Harb Perspect Biol* 6, a022616.
263. Stenmark, H. (2009). Rab GTPases as coordinators of vesicle traffic. *Nat Rev Mol Cell Biol* 10, 513-525.
264. Zhen, Y., and Stenmark, H. (2015). Cellular functions of Rab GTPases at a glance. *J Cell Sci* 128, 3171-3176.
265. Hutagalung, A.H., and Novick, P.J. (2011). Role of Rab GTPases in membrane traffic and cell physiology. *Physiol Rev* 91, 119-149.
266. Cherfils, J., and Zeghouf, M. (2013). Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* 93, 269-309.
267. Eathiraj, S., Pan, X., Ritacco, C., and Lambright, D.G. (2005). Structural basis of family-wide Rab GTPase recognition by rabenosyn-5. *Nat New Biol* 436, 415-419.
268. D'Adamo, P., Masetti, M., Bianchi, V., More, L., Mignogna, M.L., Giannandrea, M., and Gatti, S. (2014). RAB GTPases and RAB-interacting proteins and their role in the control of cognitive functions. *Neurosci Biobehav Rev* 46 Pt 2, 302-314.
269. Scapin, S.M., Carneiro, F.R., Alves, A.C., Medrano, F.J., Guimaraes, B.G., and Zanchin, N.I. (2006). The crystal structure of the small GTPase Rab11b reveals critical differences relative to the Rab11a isoform. *J Struct*

- Biol 154, 260-268.
270. Krieger, E., and Vriend, G. (2015). New ways to boost molecular dynamics simulations. *J Comput Chem* 36, 996-1007.
271. Silvis, M.R., Bertrand, C.A., Ameen, N., Golin-Bisello, F., Butterworth, M.B., Frizzell, R.A., and Bradbury, N.A. (2009). Rab11b regulates the apical recycling of the cystic fibrosis transmembrane conductance regulator in polarized intestinal epithelial cells. *Mol Biol Cell* 20, 2337-2350.
272. Ullrich, O., Reinsch, S., Urbe, S., Zerial, M., and Parton, R.G. (1996). Rab11 regulates recycling through the pericentriolar recycling endosome. *J Cell Biol* 135, 913-924.
273. Schlierf, B., Fey, G.H., Hauber, J., Hocke, G.M., and Rosorius, O. (2000). Rab11b is essential for recycling of transferrin to the plasma membrane. *Exp Cell Res* 259, 257-265.
274. Sprang, S.R., and Coleman, D.E. (1998). Invasion of the nucleotide snatchers: structural insights into the mechanism of G protein GEFs. *Cell* 95, 155-158.
275. Dumas, J.J., Zhu, Z., Connolly, J.L., and Lambright, D.G. (1999). Structural basis of activation and GTP hydrolysis in Rab proteins. *Structure* 7, 413-423.
276. Welz, T., Wellbourne-Wood, J., and Kerkhoff, E. (2014). Orchestration of cell surface proteins by Rab11. *Trends Cell Biol* 24, 407-415.
277. Kelly, E.E., Horgan, C.P., and McCaffrey, M.W. (2012). Rab11 proteins in health and disease. *Biochem Soc Trans* 40, 1360-1367.
278. Goldenring, J.R., Shen, K.R., Vaughan, H.D., and Modlin, I.M. (1993). Identification of a small GTP-binding protein, Rab25, expressed in the gastrointestinal mucosa, kidney, and lung. *J Biol Chem* 268, 18419-18422.
279. Lai, F., Stubbs, L., and Artzt, K. (1994). Molecular analysis of mouse Rab11b: a new type of mammalian YPT/Rab protein. *Genomics* 22, 610-616.
280. Wilcke, M., Johannes, L., Galli, T., Mayau, V., Goud, B., and Salamero, J. (2000). Rab11 regulates the compartmentalization of early endosomes required for efficient transport from early endosomes to the trans-golgi network. *J Cell Biol* 151, 1207-1220.
281. Knodler, A., Feng, S., Zhang, J., Zhang, X., Das, A., Peranen, J., and Guo, W. (2010). Coordination of Rab8 and Rab11 in primary ciliogenesis. *Proc Natl Acad Sci U S A* 107, 6346-6351.
282. Westlake, C.J., Baye, L.M., Nachury, M.V., Wright, K.J., Ervin, K.E., Phu, L., Chalouni, C., Beck, J.S., Kirkpatrick, D.S., Slusarski, D.C., et al. (2011). Primary cilia membrane assembly is initiated by Rab11 and transport protein particle II (TRAPP II) complex-dependent trafficking of Rabin8 to the centrosome. *Proc Natl Acad Sci U S A* 108, 2759-2764.
283. Feng, S., Knodler, A., Ren, J., Zhang, J., Zhang, X., Hong, Y., Huang, S., Peranen, J., and Guo, W. (2012). A Rab8 guanine nucleotide exchange factor-effector interaction network regulates primary ciliogenesis. *J Biol Chem* 287, 15602-15609.
284. Hehnl, H., Chen, C.T., Powers, C.M., Liu, H.L., and Doxsey, S. (2012). The centrosome regulates the Rab11-dependent recycling endosome pathway at appendages of the mother centriole. *Curr Biol* 22, 1944-1950.
285. Letteboer, S.J., and Roepman, R. (2008). Versatile screening for binary protein-protein interactions by yeast two-hybrid mating. *Methods Mol Biol* 484, 145-159.
286. Houge, G., Rasmussen, I.H., and Hovland, R. (2012). Loss-of-Function CNKSR2 Mutation Is a Likely Cause of Non-Syndromic X-Linked Intellectual Disability. *Mol Syndromol* 2, 60-63.

287. Damiano, J.A., Burgess, R., Kivity, S., Lerman-Sagie, T., Afawi, Z., Scheffer, I.E., Berkovic, S.F., and Hildebrand, M.S. (2017). Frequency of CNKSR2 mutation in the X-linked epilepsy-aphasia spectrum. *Epilepsia* 58, e40-e43.
288. Vaags, A.K., Bowdin, S., Smith, M.L., Gilbert-Dussardier, B., Brocke-Holmefjord, K.S., Sinopoli, K., Gilles, C., Haaland, T.B., Vincent-Delorme, C., Lagrue, E., et al. (2014). Absent CNKSR2 causes seizures and intellectual, attention, and language deficits. *Ann Neurol* 76, 758-764.
289. Yoo, H.Y., Sung, M.K., Lee, S.H., Kim, S., Lee, H., Park, S., Kim, S.C., Lee, B., Rho, K., Lee, J.E., et al. (2014). A recurrent inactivating mutation in RHOA GTPase in angioimmunoblastic T cell lymphoma. *Nat Genet* 46, 371-375.
290. Palomero, T., Couronne, L., Khiabanian, H., Kim, M.Y., Ambesi-Impiombato, A., Perez-Garcia, A., Carpenter, Z., Abate, F., Allegretta, M., Haydu, J.E., et al. (2014). Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nat Genet* 46, 166-170.
291. Sakata-Yanagimoto, M., Enami, T., Yoshida, K., Shiraishi, Y., Ishii, R., Miyake, Y., Muto, H., Tsuyama, N., Sato-Otsubo, A., Okuno, Y., et al. (2014). Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat Genet* 46, 171-175.
292. Feig, L.A. (1999). Tools of the trade: use of dominant-inhibitory mutants of Ras-family GTPases. *Nat Cell Biol* 1, E25-27.
293. Barrowman, J., Bhandari, D., Reinisch, K., and Ferro-Novick, S. (2010). TRAPP complexes in membrane traffic: convergence through a common Rab. *Nat Rev Mol Cell Biol* 11, 759-763.
294. Morozova, N., Liang, Y., Tokarev, A.A., Chen, S.H., Cox, R., Andrejic, J., Lipatova, Z., Sciorra, V.A., Emr, S.D., and Segev, N. (2006). TRAPPII subunits are required for the specificity switch of a Ypt-Rab GEF. *Nat Cell Biol* 8, 1263-1269.
295. Jones, S., Newman, C., Liu, F., and Segev, N. (2000). The TRAPP complex is a nucleotide exchanger for Ypt1 and Ypt31/32. *Mol Biol Cell* 11, 4403-4411.
296. Thomas, L.L., and Fromme, J.C. (2016). GTPase cross talk regulates TRAPPII activation of Rab11 homologues during vesicle biogenesis. *J Cell Biol* 215, 499-513.
297. Sakaguchi, A., Sato, M., Sato, K., Gengyo-Ando, K., Yorimitsu, T., Nakai, J., Hara, T., Sato, K., and Sato, K. (2015). REI-1 Is a Guanine Nucleotide Exchange Factor Regulating RAB-11 Localization and Function in *C. elegans* Embryos. *Dev Cell* 35, 211-221.
298. Chiba, S., Amagai, Y., Homma, Y., Fukuda, M., and Mizuno, K. (2013). NDR2-mediated Rabin8 phosphorylation is crucial for ciliogenesis by switching binding specificity from phosphatidylserine to Sec15. *EMBO J* 32, 874-885.
299. Roosing, S., Lamers, I.J., de Vrieze, E., van den Born, L.I., Lambertus, S., Arts, H.H., Group, P.B.S., Peters, T.A., Hoyng, C.B., Kremer, H., et al. (2014). Disruption of the basal body protein POC1B results in autosomal-recessive cone-rod dystrophy. *Am J Hum Genet* 95, 131-142.
300. Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* 224, 213-232.
301. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676-682.

302. Schindelin, J., Rueden, C.T., Hiner, M.C., and Eliceiri, K.W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 82, 518-529.
303. Lubs, H.A., Stevenson, R.E., and Schwartz, C.E. (2012). Fragile X and X-linked intellectual disability: four decades of discovery. *Am J Hum Genet* 90, 579-590.
304. Piton, A., Redin, C., and Mandel, J.L. (2013). XLID-causing mutations and associated genes challenged in light of data from large-scale human exome sequencing. *Am J Hum Genet* 93, 368-383.
305. Tarpey, P.S., Smith, R., Pleasance, E., Whibley, A., Edkins, S., Hardy, C., O'Meara, S., Latimer, C., Dicks, E., Menzies, A., et al. (2009). A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 41, 535-543.
306. Hu, H., Haas, S.A., Chelly, J., Van Esch, H., Raynaud, M., de Brouwer, A.P., Weinert, S., Froyen, G., Frints, S.G., Laumonnier, F., et al. (2015). X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. *Mol Psychiatry*.
307. Whibley, A.C., Plagnol, V., Tarpey, P.S., Abidi, F., Fullston, T., Choma, M.K., Boucher, C.A., Shepherd, L., Willatt, L., Parkin, G., et al. (2010). Fine-scale survey of X chromosome copy number variants and indels underlying intellectual disability. *Am J Hum Genet* 87, 173-188.
308. Stevens-Kroef, M.J., van den Berg, E., Olde Weghuis, D., Geurts van Kessel, A., Pfundt, R., Linssen-Wiersma, M., Benjamins, M., Dijkhuizen, T., Groenen, P.J., and Simons, A. (2014). Identification of prognostic relevant chromosomal abnormalities in chronic lymphocytic leukemia using microarray-based genomic profiling. *Mol Cytogenet* 7, 3.
309. Brett, M., McPherson, J., Zang, Z.J., Lai, A., Tan, E.S., Ng, I., Ong, L.C., Cham, B., Tan, P., Rozen, S., et al. (2014). Massively parallel sequencing of patients with intellectual disability, congenital anomalies and/or autism spectrum disorders with a targeted gene panel. *PLoS One* 9, e93409.
310. Wiszniewska, J., Bi, W., Shaw, C., Stankiewicz, P., Kang, S.H., Pursley, A.N., Lalani, S., Hixson, P., Gambin, T., Tsai, C.H., et al. (2014). Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing. *Eur J Hum Genet* 22, 79-87.
311. Lepri, F.R., Scavelli, R., Digilio, M.C., Gnazzo, M., Grotta, S., Dentici, M.L., Pisaneschi, E., Sirlito, P., Capolino, R., Baban, A., et al. (2014). Diagnosis of Noonan syndrome and related disorders using target next generation sequencing. *BMC Med Genet* 15, 14.
312. Faesen, A.C., Luna-Vargas, M.P., and Sixma, T.K. (2012). The role of UBL domains in ubiquitin-specific proteases. *Biochem Soc Trans* 40, 539-545.
313. Murtaza, M., Jolly, L.A., Gecz, J., and Wood, S.A. (2015). La FAM fatale: USP9X in development and disease. *Cell Mol Life Sci* 72, 2075-2089.
314. Khut, P.Y., Tucker, B., Lardelli, M., and Wood, S.A. (2007). Evolutionary and expression analysis of the zebrafish deubiquitylating enzyme, *usp9*. *Zebrafish* 4, 95-101.
315. Jolly, L.A., Taylor, V., and Wood, S.A. (2009). USP9X enhances the polarity and self-renewal of embryonic stem cell-derived neural progenitors. *Mol Biol Cell* 20, 2015-2029.
316. Wood, S.A., Pascoe, W.S., Ru, K., Yamada, T., Hirschman, J., Kemler, R., and Mattick, J.S. (1997). Cloning and expression analysis of a novel mouse gene with sequence similarity to the *Drosophila fat facets* gene. *Mech Dev* 63, 29-38.
317. Jones, M.H., Furlong, R.A., Burkin, H., Chalmers, I.J., Brown, G.M., Khwaja, O., and Affara, N.A. (1996). The Dro-

- sophila developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. *Hum Mol Genet* 5, 1695-1701.
318. Fischer-Vize, J.A., Rubin, G.M., and Lehmann, R. (1992). The fat facets gene is required for Drosophila eye and embryo development. *Development* 116, 985-1000.
 319. Dupont, S., Mamidi, A., Cordenonsi, M., Montagner, M., Zacchigna, L., Adorno, M., Martello, G., Stinchfield, M.J., Soligo, S., Morsut, L., et al. (2009). FAM/USP9x, a deubiquitinating enzyme essential for TGFbeta signaling, controls Smad4 monoubiquitination. *Cell* 136, 123-135.
 320. Friocourt, G., Kappeler, C., Saillour, Y., Fauchereau, F., Rodriguez, M.S., Bahi, N., Vinet, M.C., Chafey, P., Poirier, K., Taya, S., et al. (2005). Doublecortin interacts with the ubiquitin protease DFFRX, which associates with microtubules in neuronal processes. *Mol Cell Neurosci* 28, 153-164.
 321. Xie, Y., Avello, M., Schirle, M., McWhinnie, E., Feng, Y., Bric-Furlong, E., Wilson, C., Nathans, R., Zhang, J., Kirschner, M.W., et al. (2013). Deubiquitinase FAM/USP9X interacts with the E3 ubiquitin ligase SMURF1 protein and protects it from ligase activity-dependent self-degradation. *J Biol Chem* 288, 2976-2985.
 322. Agrawal, P., Chen, Y.T., Schilling, B., Gibson, B.W., and Hughes, R.E. (2012). Ubiquitin-specific peptidase 9, X-linked (USP9X) modulates activity of mammalian target of rapamycin (mTOR). *J Biol Chem* 287, 21164-21175.
 323. Taya, S., Yamamoto, T., Kanai-Azuma, M., Wood, S.A., and Kaibuchi, K. (1999). The deubiquitinating enzyme Fam interacts with and stabilizes beta-catenin. *Genes Cells* 4, 757-767.
 324. Taya, S., Yamamoto, T., Kano, K., Kawano, Y., Iwamatsu, A., Tsuchiya, T., Tanaka, K., Kanai-Azuma, M., Wood, S.A., Mattick, J.S., et al. (1998). The Ras target AF-6 is a substrate of the fam deubiquitinating enzyme. *J Cell Biol* 142, 1053-1062.
 325. Mouchantaf, R., Azakir, B.A., McPherson, P.S., Millard, S.M., Wood, S.A., and Angers, A. (2006). The ubiquitin ligase itch is auto-ubiquitylated in vivo and in vitro but is protected from degradation by interacting with the deubiquitylating enzyme FAM/USP9X. *J Biol Chem* 281, 38738-38747.
 326. Al-Hakim, A.K., Zagorska, A., Chapman, L., Deak, M., Pegg, M., and Alessi, D.R. (2008). Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem J* 411, 249-260.
 327. Luise, C., Capra, M., Donzelli, M., Mazzarol, G., Jodice, M.G., Nuciforo, P., Viale, G., Di Fiore, P.P., and Confalonieri, S. (2011). An atlas of altered expression of deubiquitinating enzymes in human cancer. *PLoS One* 6, e15891.
 328. Schwickart, M., Huang, X., Lill, J.R., Liu, J., Ferrando, R., French, D.M., Maecker, H., O'Rourke, K., Bazan, F., Eastham-Anderson, J., et al. (2010). Deubiquitinase USP9X stabilizes MCL1 and promotes tumour cell survival. *Nat New Biol* 463, 103-107.
 329. Deng, X., Berletch, J.B., Nguyen, D.K., and Distech, C.M. (2014). X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet* 15, 367-378.
 330. Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S., and Brown, C.J. (2015). Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Hum Mol Genet* 24, 1528-1539.
 331. Berletch, J.B., Yang, F., Xu, J., Carrel, L., and Distech, C.M. (2011). Genes that escape from X inactivation. *Hum Genet* 130, 237-245.
 332. Yang, C., Chapman, A.G., Kelsey, A.D., Minks, J., Cotton, A.M., and Brown, C.J. (2011). X-chromosome inactivation: molecular mechanisms from the human perspective. *Hum Genet* 130, 175-185.

333. Morleo, M., and Franco, B. (2008). Dosage compensation of the mammalian X chromosome influences the phenotypic variability of X-linked dominant male-lethal disorders. *J Med Genet* 45, 401-408.
334. Carrel, L., and Willard, H.F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nat New Biol* 434, 400-404.
335. Talebizadeh, Z., Simon, S.D., and Butler, M.G. (2006). X chromosome gene expression in human tissues: male and female comparisons. *Genomics* 88, 675-681.
336. Zweier, C., Kraus, C., Brueton, L., Cole, T., Degenhardt, F., Engels, H., Gillissen-Kaesbach, G., Graul-Neumann, L., Horn, D., Hoyer, J., et al. (2013). A new face of Borjeson-Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype. *J Med Genet* 50, 838-847.
337. Peeters, S.B., Cotton, A.M., and Brown, C.J. (2014). Variable escape from X-chromosome inactivation: Identifying factors that tip the scales towards expression. *Bioessays* 36, 746-756.
338. Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M.R., Venselaar, H., Helsmoortel, C., Cho, M.T., Hoischen, A., et al. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet* 97, 343-352.
339. Xu J, T.S., Kaibuchi K, Arnold AP. (2005). Sexually dimorphic expression of Usp9x is related to sex chromosome complement in adult mouse brain. *Eur J Neurosci* 21, 3017-3022.
340. Homan, C.C., Kumar, R., Nguyen, L.S., Haan, E., Raymond, F.L., Abidi, F., Raynaud, M., Schwartz, C.E., Wood, S.A., Gecz, J., et al. (2014). Mutations in USP9X are associated with X-linked intellectual disability and disrupt neuronal cell migration and growth. *Am J Hum Genet* 94, 470-478.
341. Paemka, L., Mahajan, V.B., Ehaideb, S.N., Skeie, J.M., Tan, M.C., Wu, S., Cox, A.J., Sowers, L.P., Gecz, J., Jolly, L., et al. (2015). Seizures Are Regulated by Ubiquitin-specific Peptidase 9 X-linked (USP9X), a De-Ubiquitinase. *PLoS Genet* 11, e1005022.
342. Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* 23, 198-199.
343. Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* 100, 189-192.
344. Stegeman, S., Jolly, L.A., Premaratne, S., Gecz, J., Richards, L.J., Mackay-Sim, A., and Wood, S.A. (2013). Loss of Usp9x disrupts cortical architecture, hippocampal development and TGFbeta-mediated axonogenesis. *PLoS One* 8, e68287.
345. Ware, S.M., Aygun, M.G., and Hildebrandt, F. (2011). Spectrum of clinical diseases caused by disorders of primary cilia. *Proc Am Thorac Soc* 8, 444-450.
346. Tucker, R.W., Pardee, A.B., and Fujiwara, K. (1979). Centriole ciliation is related to quiescence and DNA synthesis in 3T3 cells. *Cell* 17, 527-535.
347. Vorobjev, I.A., and Chentsov Yu, S. (1982). Centrioles in the cell cycle. I. Epithelial cells. *The Journal of cell biology* 93, 938-949.
348. Murray, R.Z., Jolly, L.A., and Wood, S.A. (2004). The FAM deubiquitylating enzyme localizes to multiple points of protein trafficking in epithelia, where it associates with E-cadherin and beta-catenin. *Mol Biol Cell* 15, 1591-1599.
349. Schuurs-Hoeijmakers, J.H., Oh, E.C., Vissers, L.E., Swinkels, M.E., Gilissen, C., Willemsen, M.A., Holvoet, M., Steehouwer, M., Veltman, J.A., de Vries, B.B., et al. (2012). Recurrent de novo mutations in PACS1 cause

- defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am J Hum Genet* 91, 1122-1127.
350. Willemsen, M.H., Vissers, L.E., Willemsen, M.A., van Bon, B.W., Kroes, T., de Ligt, J., de Vries, B.B., Schoots, J., Lugtenberg, D., Hamel, B.C., et al. (2012). Mutations in *DYNC1H1* cause severe intellectual disability with neuronal migration defects. *J Med Genet* 49, 179-183.
 351. Grozeva, D., Carss, K., Spasic-Boskovic, O., Parker, M.J., Archer, H., Firth, H.V., Park, S.M., Canham, N., Holder, S.E., Wilson, M., et al. (2014). De novo loss-of-function mutations in *SETD5*, encoding a methyltransferase in a 3p25 microdeletion syndrome critical region, cause intellectual disability. *Am J Hum Genet* 94, 618-624.
 352. O'Roak, B.J., Vives, L., Fu, W., Egerton, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K., et al. (2012). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* (80-) 338, 1619-1622.
 353. Vulto-van Silfhout, A.T., Rajamanickam, S., Jensik, P.J., Vergult, S., de Roker, N., Newhall, K.J., Raghavan, R., Reardon, S.N., Jarrett, K., McIntyre, T., et al. (2014). Mutations affecting the SAND domain of *DEAF1* cause intellectual disability with severe speech impairment and behavioral problems. *Am J Hum Genet* 94, 649-661.
 354. Willemsen, M.H., Nijhof, B., Fenckova, M., Nillesen, W.M., Bongers, E.M., Castells-Nobau, A., Asztalos, L., Viraigh, E., van Bon, B.W., Tezel, E., et al. (2013). *GATAD2B* loss-of-function mutations cause a recognisable syndrome with intellectual disability and are associated with learning deficits and synaptic undergrowth in *Drosophila*. *J Med Genet* 50, 507-514.
 355. Hamdan, F.F., Srouf, M., Capo-Chichi, J.M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O., et al. (2014). De novo mutations in moderate or severe intellectual disability. *PLoS Genet* 10, e1004772.
 356. Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R., and Ohara, O. (2001). Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 8, 85-95.
 357. Zhang, F., and Yu, X. (2011). WAC, a functional partner of RNF20/40, regulates histone H2B ubiquitination and gene transcription. *Mol Cell* 41, 384-397.
 358. Xu, G.M., and Arnaout, M.A. (2002). WAC, a novel WW domain-containing adapter with a coiled-coil region, is colocalized with splicing factor SC35. *Genomics* 79, 87-94.
 359. Wentzel, C., Rajcan-Separovic, E., Ruivenkamp, C.A., Chantot-Bastaraud, S., Metay, C., Andrieux, J., Anneren, G., Gijssbers, A.C., Druart, L., Hyon, C., et al. (2011). Genomic and clinical characteristics of six patients with partially overlapping interstitial deletions at 10p12p11. *Eur J Hum Genet* 19, 959-964.
 360. Okamoto, N., Hayashi, S., Masui, A., Kosaki, R., Oguri, I., Hasegawa, T., Imoto, I., Makita, Y., Hata, A., Moriyama, K., et al. (2012). Deletion at chromosome 10p11.23-p12.1 defines characteristic phenotypes with marked midface retrusion. *J Hum Genet* 57, 191-196.
 361. Shahdadpuri, R., de Vries, B., Pfundt, R., de Leeuw, N., and Reardon, W. (2008). Pseudoarthrosis of the clavicle and copper beaten skull associated with chromosome 10p11.21p12.1 microdeletion. *Am J Med Genet A* 146A, 233-237.
 362. Mroczkowski, H.J., Arnold, G., Schneck, F.X., Rajkovic, A., and Yatsenko, S.A. (2014). Interstitial 10p11.23-p12.1 microdeletions associated with developmental delay, craniofacial abnormalities, and cryptorchidism. *Am J*

- Med Genet A 164A, 2623-2626.
363. Exome Aggregation Consortium (ExAC). Cambridge, MA, USA, 2015.
 364. Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nat New Biol* 471, 473-479.
 365. modMine. Available at <http://www.modencode.org/Celniker.shtm>.
 366. Wilson, D.A., and Linster, C. (2008). Neurobiology of a simple memory. *J Neurophysiol* 100, 2-7.
 367. Typlt, M., Mirkowski, M., Azzopardi, E., Ruettiger, L., Ruth, P., and Schmid, S. (2013). Mice with deficient BK channel function show impaired prepulse inhibition and spatial learning, but normal working and spatial reference memory. *PLoS One* 8, e81270.
 368. van Bon, B.W., Oortveld, M.A., Nijtmans, L.G., Fencikova, M., Nijhof, B., Besseling, J., Vos, M., Kramer, J.M., de Leeuw, N., Castells-Nobau, A., et al. (2013). CEP89 is required for mitochondrial metabolism and neuronal function in man and fly. *Hum Mol Genet* 22, 3138-3151.
 369. Kramer, J.M., Kochinke, K., Oortveld, M.A., Marks, H., Kramer, D., de Jong, E.K., Asztalos, Z., Westwood, J.T., Stunnenberg, H.G., Sokolowski, M.B., et al. (2011). Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *PLoS Biol* 9, e1000569.
 370. vienna *Drosophila* RNAi Center. Available at www.vdrc.at.
 371. Slager, R.E., Newton, T.L., Vlangos, C.N., Finucane, B., and Elsea, S.H. (2003). Mutations in RAI1 associated with Smith-Magenis syndrome. *Nat Genet* 33, 466-468.
 372. Elsea, S.H., and Girirajan, S. (2008). Smith-Magenis syndrome. *Eur J Hum Genet* 16, 412-421.
 373. Luxan, G., Casanova, J.C., Martinez-Poveda, B., Prados, B., D'Amato, G., MacGrogan, D., Gonzalez-Rajal, A., Dobarro, D., Torroja, C., Martinez, F., et al. (2013). Mutations in the NOTCH pathway regulator MIB1 cause left ventricular noncompaction cardiomyopathy. *Nat Med* 19, 193-201.
 374. Haddad, D.M., Vilain, S., Vos, M., Esposito, G., Matta, S., Kalscheuer, V.M., Craessaerts, K., Leyssen, M., Nascimento, R.M., Vianna-Morgante, A.M., et al. (2013). Mutations in the intellectual disability gene *Ube2a* cause neuronal dysfunction and impair parkin-dependent mitophagy. *Mol Cell* 50, 831-843.
 375. ECARUCA. Available at <http://umcecaruca01.extern.umcn.nl:8080/ecaruca/ecaruca.jsp>.
 376. DECIPHER. Available at <https://decipher.sanger.ac.uk/index>.
 377. Dietzl, G., Chen, D., Schnorrer, F., Su, K.C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblaue, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nat New Biol* 448, 151-156.
 378. Bloomington *Drosophila* Stock Center. Available at www.flystocks.bio.indiana.edu.
 379. Team RC R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2014.
 380. Team, R.C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
 381. Shimojima, K., Isidor, B., Le Caignec, C., Kondo, A., Sakata, S., Ohno, K., and Yamamoto, T. (2011). A new microdeletion syndrome of 5q31.3 characterized by severe developmental delays, distinctive facial features, and delayed myelination. *Am J Med Genet A* 155A, 732-736.
 382. Hosoki, K., Ohta, T., Natsume, J., Imai, S., Okumura, A., Matsui, T., Harada, N., Bacino, C.A., Scaglia, F., Jones,

- J.Y., et al. (2012). Clinical phenotype and candidate genes for the 5q31.3 microdeletion syndrome. *Am J Med Genet A* 158A, 1891-1896.
383. Rosenfeld, J.A., Drautz, J.M., Clericuzio, C.L., Cushing, T., Raskin, S., Martin, J., Tervo, R.C., Pitarque, J.A., Nowak, D.M., Karolak, J.A., et al. (2011). Deletions and duplications of developmental pathway genes in 5q31 contribute to abnormal phenotypes. *Am J Med Genet A* 155A, 1906-1916.
384. Brown, N., Burgess, T., Forbes, R., McGillivray, G., Kornberg, A., Mandelstam, S., and Stark, Z. (2013). 5q31.3 Microdeletion syndrome: clinical and molecular characterization of two further cases. *Am J Med Genet A* 161A, 2604-2608.
385. Bonaglia, M.C., Zanutta, N., Giorda, R., D'Angelo, G., and Zucca, C. (2015). Long-term follow-up of a patient with 5q31.3 microdeletion syndrome and the smallest de novo 5q31.2q31.3 deletion involving PURA. *Mol Cytogenet* 8, 89.
386. Lalani, S.R., Zhang, J., Schaaf, C.P., Brown, C.W., Magoulas, P., Tsai, A.C., El-Gharbawy, A., Wierenga, K.J., Bartholomew, D., Fong, C.T., et al. (2014). Mutations in PURA cause profound neonatal hypotonia, seizures, and encephalopathy in 5q31.3 microdeletion syndrome. *Am J Hum Genet* 95, 579-583.
387. Hunt, D., Leventer, R.J., Simons, C., Taft, R., Swoboda, K.J., Gawne-Cain, M., study, D.D.D., Magee, A.C., Turnpenny, P.D., and Baralle, D. (2014). Whole exome sequencing in family trios reveals de novo mutations in PURA as a cause of severe neurodevelopmental delay and learning disability. *J Med Genet* 51, 806-813.
388. Tanaka, A.J., Bai, R., Cho, M.T., Anyane-Yeboah, K., Ahimaz, P., Wilson, A.L., Kendall, F., Hay, B., Moss, T., Nardini, M., et al. (2015). De novo mutations in PURA are associated with hypotonia and developmental delay. *Cold Spring Harbor molecular case studies* 1, a000356.
389. Okamoto, N., Nakao, H., Niihori, T., and Aoki, Y. (2017). A patient with a novel purine-rich element binding protein A (PURA) mutation. *Congenit Anom (Kyoto)*.
390. Rezkalla, J., Von Wald, T., and Hansen, K.A. (2017). Premature Thelarche and the PURA Syndrome. *Obstet Gynecol*.
391. Graebisch, A., Roche, S., and Niessing, D. (2009). X-ray structure of Pur-alpha reveals a Whirly-like fold and an unusual nucleic-acid binding surface. *Proc Natl Acad Sci U S A* 106, 18521-18526.
392. Weber, J., Bao, H., Hartmuller, C., Wang, Z., Windhager, A., Janowski, R., Madl, T., Jin, P., and Niessing, D. (2016). Structural basis of nucleic-acid recognition and double-strand unwinding by the essential neuronal protein Pur-alpha. *eLife* 5.
393. White, M.K., Johnson, E.M., and Khalili, K. (2009). Multiple roles for Puralpha in cellular and viral regulation. *Cell Cycle* 8, 1-7.
394. Yuan C, L.P., Guo S, Zhang B, Sun T, Cui J. (2016). The Role of Pura in Neuronal Development, the Progress in the Current Researches. *J Neurol Neurosci* 7.
395. Hokkanen, S., Feldmann, H.M., Ding, H., Jung, C.K., Bojarski, L., Renner-Muller, I., Schuller, U., Kretzschmar, H., Wolf, E., and Herms, J. (2012). Lack of Pur-alpha alters postnatal brain development and causes megalencephaly. *Hum Mol Genet* 21, 473-484.
396. Khalili, K., Del Valle, L., Muralidharan, V., Gault, W.J., Darbinian, N., Otte, J., Meier, E., Johnson, E.M., Daniel, D.C., Kinoshita, Y., et al. (2003). Puralpha is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol Cell Biol* 23, 6857-6875.
397. Barbe, M.F., Krueger, J.J., Loomis, R., Otte, J., and Gordon, J. (2016). Memory deficits, gait ataxia and neuronal

- loss in the hippocampus and cerebellum in mice that are heterozygous for Pur-alpha. *Neuroscience* 337, 177-190.
398. Ansari, M., Poke, G., Ferry, Q., Williamson, K., Aldridge, R., Meynert, A.M., Bengani, H., Chan, C.Y., Kayserili, H., Avci, S., et al. (2014). Genetic heterogeneity in Cornelia de Lange syndrome (CdLS) and CdLS-like phenotypes with observed and predicted levels of mosaicism. *J Med Genet* 51, 659-668.
 399. Ferry, Q., Steinberg, J., Webber, C., FitzPatrick, D.R., Ponting, C.P., Zisserman, A., and Nellaker, C. (2014). Diagnostically relevant facial gestalt information from ordinary photos. *eLife* 3, e02020.
 400. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10, 845-858.
 401. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501.
 402. Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., et al. (2010). Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 42, 30-35.
 403. Jansen, S., Geuer, S., Pfundt, R., Brough, R., Ghongane, P., Herkert, J.C., Marco, E.J., Willemsen, M.H., Kleefstra, T., Hannibal, M., et al. (2017). De Novo Truncating Mutations in the Last and Penultimate Exons of PPM1D Cause an Intellectual Disability Syndrome. *Am J Hum Genet* 100, 650-658.
 404. Schafgen, J., Cremer, K., Becker, J., Wieland, T., Zink, A.M., Kim, S., Windheuser, I.C., Kreiss, M., Aretz, S., Strom, T.M., et al. (2016). De novo nonsense and frameshift variants of TCF20 in individuals with intellectual disability and postnatal overgrowth. *Eur J Hum Genet* 24, 1739-1745.
 405. Bramswig, N.C., Ludecke, H.J., Pettersson, M., Albrecht, B., Bernier, R.A., Cremer, K., Eichler, E.E., Falkenstein, D., Gerdts, J., Jansen, S., et al. (2017). Identification of new TRIP12 variants and detailed clinical evaluation of individuals with non-syndromic intellectual disability with or without autism. *Hum Genet* 136, 179-192.
 406. Zhang, J., Gambin, T., Yuan, B., Szafranski, P., Rosenfeld, J.A., Balwi, M.A., Alswaid, A., Al-Gazali, L., Shamsi, A.M.A., Komara, M., et al. (2017). Haploinsufficiency of the E3 ubiquitin-protein ligase gene TRIP12 causes intellectual disability with or without autism spectrum disorders, speech delay, and dysmorphic features. *Hum Genet* 136, 377-386.
 407. Au, P.Y., Huang, L., Broley, S., Gallagher, L., Creede, E., Lahey, D., Ordorica, S., Mina, K., Boycott, K.M., Baynam, G., et al. (2017). Two females with mutations in USP9X highlight the variable expressivity of the intellectual disability syndrome. *Eur J Med Genet*.
 408. Tokita, M.J., Braxton, A.A., Shao, Y., Lewis, A.M., Vincent, M., Kury, S., Besnard, T., Isidor, B., Latypova, X., Bezieau, S., et al. (2016). De Novo Truncating Variants in SON Cause Intellectual Disability, Congenital Malformations, and Failure to Thrive. *Am J Hum Genet* 99, 720-727.
 409. Nesbitt, A., Bhoj, E.J., McDonald Gibson, K., Yu, Z., Denenberg, E., Sarmady, M., Tischler, T., Cao, K., Dubbs, H., Zackai, E.H., et al. (2015). Exome sequencing expands the mechanism of SOX5-associated intellectual disability: A case presentation with review of sox-related disorders. *Am J Med Genet A* 167A, 2548-2554.
 410. Babbs, C., Lloyd, D., Pagnamenta, A.T., Twigg, S.R., Green, J., McGowan, S.J., Mirza, G., Naples, R., Sharma, V.P., Volpi, E.V., et al. (2014). De novo and rare inherited mutations implicate the transcriptional coregulator TCF20/SPBP in autism spectrum disorder. *J Med Genet* 51, 737-747.
 411. DeSanto, C., D'Aco, K., Araujo, G.C., Shannon, N., Study, D.D.D., Vernon, H., Rahrig, A., Monaghan, K.G., Niu,

- Z., Vitazka, P., et al. (2015). WAC loss-of-function mutations cause a recognisable syndrome characterised by dysmorphic features, developmental delay and hypotonia and recapitulate 10p11.23 microdeletion syndrome. *J Med Genet* 52, 754-761.
412. Hoischen, A., Krumm, N., and Eichler, E.E. (2014). Prioritization of neurodevelopmental disease genes by discovery of new mutations. *Nat Neurosci* 17, 764-772.
413. Santen, G.W., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M., Kant, S.G., Snoeck, I.N., Peeters, E.A., Hilhorst-Hofstee, Y., et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet* 44, 379-380.
414. Sirmaci, A., Spiliopoulos, M., Brancati, F., Powell, E., Duman, D., Abrams, A., Bademci, G., Agolini, E., Guo, S., Konuk, B., et al. (2011). Mutations in ANKRD11 cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am J Hum Genet* 89, 289-294.
415. Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I., Beck, A.E., Tabor, H.K., Cooper, G.M., Mefford, H.C., et al. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 42, 790-793.
416. Hoischen, A., van Bon, B.W., Gilissen, C., Arts, P., van Lier, B., Steehouwer, M., de Vries, P., de Reuver, R., Wieskamp, N., Mortier, G., et al. (2010). De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* 42, 483-485.
417. Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nat New Biol* 485, 237-241.
418. Lehman, A., Thouta, S., Mancini, G.M.S., Naidu, S., van Slegtenhorst, M., McWalter, K., Person, R., Mwenifumbo, J., Salvarinova, R., Study, C., et al. (2017). Loss-of-Function and Gain-of-Function Mutations in KCNQ5 Cause Intellectual Disability or Epileptic Encephalopathy. *Am J Hum Genet* 101, 65-74.
419. Okur, V., Cho, M.T., Henderson, L., Retterer, K., Schneider, M., Sattler, S., Niyazov, D., Azage, M., Smith, S., Picker, J., et al. (2016). De novo mutations in CSNK2A1 are associated with neurodevelopmental abnormalities and dysmorphic features. *Hum Genet* 135, 699-705.
420. Oti, M., Snel, B., Huynen, M.A., and Brunner, H.G. (2006). Predicting disease genes using protein-protein interactions. *J Med Genet* 43, 691-698.
421. Roosing, S., Romani, M., Isrie, M., Rosti, R.O., Micalizzi, A., Musaev, D., Mazza, T., Al-Gazali, L., Altunoglu, U., Boltshauser, E., et al. (2016). Mutations in CEP120 cause Joubert syndrome as well as complex ciliopathy phenotypes. *J Med Genet* 53, 608-615.
422. Kleefstra, T., Kramer, J.M., Neveling, K., Willemsen, M.H., Koemans, T.S., Vissers, L.E., Wissink-Lindhout, W., Fenckova, M., van den Akker, W.M., Kasri, N.N., et al. (2012). Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am J Hum Genet* 91, 73-82.
423. Flex, E., Jaiswal, M., Pantaleoni, F., Martinelli, S., Strullu, M., Fansa, E.K., Caye, A., De Luca, A., Lepri, F., Dvorsky, R., et al. (2014). Activating mutations in RRAS underlie a phenotype within the RASopathy spectrum and contribute to leukaemogenesis. *Hum Mol Genet* 23, 4315-4327.
424. Robinson, P.N., and Mundlos, S. (2010). The human phenotype ontology. *Clin Genet* 77, 525-534.
425. Knight, S.J., Regan, R., Nicod, A., Horsley, S.W., Kearney, L., Homfray, T., Winter, R.M., Bolton, P., and Flint, J. (1999). Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 354,

1676-1681.

426. Keller, M.C., and Miller, G. (2006). Resolving the paradox of common, harmful, heritable mental disorders: which evolutionary genetic models work best? *Behav Brain Sci* 29, 385-404; discussion 405-352.
427. Uher, R. (2009). The role of genetic variation in the causation of mental illness: an evolution-informed framework. *Mol Psychiatry* 14, 1072-1082.
428. Bogershausen, N., Gatinois, V., Riehmer, V., Kayserili, H., Becker, J., Thoenes, M., Simsek-Kiper, P.O., Barat-Houari, M., Elcioglu, N.H., Wieczorek, D., et al. (2016). Mutation Update for Kabuki Syndrome Genes KMT2D and KDM6A and Further Delineation of X-Linked Kabuki Syndrome Subtype 2. *Hum Mutat* 37, 847-864.
429. Ockeloen, C.W., Willemsen, M.H., de Munnik, S., van Bon, B.W., de Leeuw, N., Verrips, A., Kant, S.G., Jones, E.A., Brunner, H.G., van Loon, R.L., et al. (2015). Further delineation of the KBG syndrome caused by ANKRD11 aberrations. *Eur J Hum Genet* 23, 1270.
430. Tatton-Brown, K., and Rahman, N. (1993). EZH2-Related Overgrowth. In *GeneReviews*(®), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, H.C. Mefford, K. Stephens, A. Amemiya, and N. Ledbetter, eds. (Seattle (WA)).
431. Allanson, J.E., and Roberts, A.E. (1993). Noonan Syndrome. In *GeneReviews*(®), M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H. Bean, K. Stephens, and A. Amemiya, eds. (Seattle (WA)).
432. Ramoni, R.B., Mulvihill, J.J., Adams, D.R., Allard, P., Ashley, E.A., Bernstein, J.A., Gahl, W.A., Hamid, R., Loscalzo, J., McCray, A.T., et al. (2017). The Undiagnosed Diseases Network: Accelerating Discovery about Health and Disease. *Am J Hum Genet* 100, 185-192.
433. Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *Am J Hum Genet* 97, 199-215.
434. Veltman, J.A., and Brunner, H.G. (2012). De novo mutations in human genetic disease. *Nat Rev Genet* 13, 565-575.
435. Gabriele, M., Vulto-van Silfhout, A.T., Germain, P.L., Vitriolo, A., Kumar, R., Douglas, E., Haan, E., Kosaki, K., Take-nouchi, T., Rauch, A., et al. (2017). YY1 Haploinsufficiency Causes an Intellectual Disability Syndrome Featuring Transcriptional and Chromatin Dysfunction. *Am J Hum Genet* 100, 907-925.
436. Lehalle, D., Mosca-Boidron, A.L., Begtrup, A., Boute-Benejean, O., Charles, P., Cho, M.T., Clarkson, A., Devinsky, O., Duffourd, Y., Duplomb-Jego, L., et al. (2017). STAG1 mutations cause a novel cohesinopathy characterized by unspecific syndromic intellectual disability. *J Med Genet* 54, 479-488.
437. Mignot, C., von Stulpnagel, C., Nava, C., Ville, D., Sanlaville, D., Lesca, G., Rastetter, A., Gachet, B., Marie, Y., Korenke, G.C., et al. (2016). Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet* 53, 511-522.
438. Santen, G.W., Clayton-Smith, J., and consortium, A.B.C. (2014). The ARID1B phenotype: what we have learned so far. *Am J Med Genet C Semin Med Genet* 166C, 276-289.
439. Sim, J.C., White, S.M., and Lockhart, P.J. (2015). ARID1B-mediated disorders: Mutations and possible mechanisms. *Intractable Rare Dis Res* 4, 17-23.
440. Goldenberg, A., Riccardi, F., Tessier, A., Pfundt, R., Busa, T., Cacciagli, P., Capri, Y., Coutton, C., Delahaye-Duriez, A., Frebourg, T., et al. (2016). Clinical and molecular findings in 39 patients with KBG syndrome caused by deletion or mutation of ANKRD11. *Am J Med Genet A* 170, 2847-2859.

441. Low, K., Ashraf, T., Canham, N., Clayton-Smith, J., Deshpande, C., Donaldson, A., Fisher, R., Flinter, F., Foulds, N., Fryer, A., et al. (2016). Clinical and genetic aspects of KBG syndrome. *Am J Med Genet A* 170, 2835-2846.
442. Posthuma, D., and Polderman, T.J. (2013). What have we learned from recent twin studies about the etiology of neurodevelopmental disorders? *Curr Opin Neurol* 26, 111-121.
443. Aoki, Y., Niihori, T., Kawame, H., Kurosawa, K., Ohashi, H., Tanaka, Y., Filocamo, M., Kato, K., Suzuki, Y., Kure, S., et al. (2005). Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat Genet* 37, 1038-1040.
444. Wangler, M.F., Yamamoto, S., Chao, H.T., Posey, J.E., Westerfield, M., Postlethwait, J., Members of the Undiagnosed Diseases, N., Hieter, P., Boycott, K.M., Campeau, P.M., et al. (2017). Model Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research. *Genetics* 207, 9-27.
445. van Bokhoven, H., and Brunner, H.G. (2002). Splitting p63. *Am J Hum Genet* 71, 1-13.
446. Findlay, G.M., Boyle, E.A., Hause, R.J., Klein, J.C., and Shendure, J. (2014). Saturation editing of genomic regions by multiplex homology-directed repair. *Nat New Biol* 513, 120-123.
447. Sariyer, I.K., Sariyer, R., Otte, J., and Gordon, J. (2016). Pur-Alpha Induces JCV Gene Expression and Viral Replication by Suppressing SRSF1 in Glial Cells. *PLoS One* 11, e0156819.
448. Mayorga, L., Gamboni, B., Mampel, A., and Roque, M. (2018). A frame-shift deletion in the PURA gene associates with a new clinical finding: Hypoglycorrhachia. Is GLUT1 a new PURA target? *Mol Genet Metab*.
449. Leen, W.G., Klepper, J., Verbeek, M.M., Leferink, M., Hofste, T., van Engelen, B.G., Wevers, R.A., Arthur, T., Bahi-Buisson, N., Ballhausen, D., et al. (2010). Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. *Brain* 133, 655-670.
450. Yang, H., Wang, D., Engelstad, K., Bagay, L., Wei, Y., Rotstein, M., Aggarwal, V., Levy, B., Ma, L., Chung, W.K., et al. (2011). Glut1 deficiency syndrome and erythrocyte glucose uptake assay. *Ann Neurol* 70, 996-1005.
451. Seidner, G., Alvarez, M.G., Yeh, J.L., O'Driscoll, K.R., Klepper, J., Stump, T.S., Wang, D., Spinner, N.B., Birnbaum, M.J., and De Vivo, D.C. (1998). GLUT-1 deficiency syndrome caused by haploinsufficiency of the blood-brain barrier hexose carrier. *Nat Genet* 18, 188-191.
452. Bem, D., Yoshimura, S., Nunes-Bastos, R., Bond, F.C., Kurian, M.A., Rahman, F., Handley, M.T., Hadzhiev, Y., Masood, I., Straatman-Iwanowska, A.A., et al. (2011). Loss-of-function mutations in RAB18 cause Warburg micro syndrome. *Am J Hum Genet* 88, 499-507.
453. Schubbert, S., Zenker, M., Rowe, S.L., Boll, S., Klein, C., Bollag, G., van der Burgt, I., Musante, L., Kalscheuer, V., Wehner, L.E., et al. (2006). Germline KRAS mutations cause Noonan syndrome. *Nat Genet* 38, 331-336.
454. Cirstea, I.C., Kutsche, K., Dvorsky, R., Gremer, L., Carta, C., Horn, D., Roberts, A.E., Lepri, F., Merbitz-Zahradnik, T., Konig, R., et al. (2010). A restricted spectrum of NRAS mutations causes Noonan syndrome. *Nat Genet* 42, 27-29.
455. Aoki, Y., Niihori, T., Banjo, T., Okamoto, N., Mizuno, S., Kurosawa, K., Ogata, T., Takada, F., Yano, M., Ando, T., et al. (2013). Gain-of-function mutations in RIT1 cause Noonan syndrome, a RAS/MAPK pathway syndrome. *Am J Hum Genet* 93, 173-180.
456. Niihori, T., Aoki, Y., Narumi, Y., Neri, G., Cave, H., Verloes, A., Okamoto, N., Hennekam, R.C., Gillesen-Kaesbach, G., Wiczorek, D., et al. (2006). Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome. *Nat Genet* 38, 294-296.
457. Kratz, C.P., Zampino, G., Kriek, M., Kant, S.G., Leoni, C., Pantaleoni, F., Oudesluys-Murphy, A.M., Di Rocco, C.,

- Kloska, S.P., Tartaglia, M., et al. (2009). Craniosynostosis in patients with Noonan syndrome caused by germline KRAS mutations. *Am J Med Genet A* 149A, 1036-1040.
458. Carta, C., Pantaleoni, F., Bocchinfuso, G., Stella, L., Vasta, I., Sarkozy, A., Digilio, C., Palleschi, A., Pizzuti, A., Grammatico, P., et al. (2006). Germline missense mutations affecting KRAS Isoform B are associated with a severe Noonan syndrome phenotype. *Am J Hum Genet* 79, 129-135.
459. Zenker, M., Lehmann, K., Schulz, A.L., Barth, H., Hansmann, D., Koenig, R., Korinthenberg, R., Kreiss-Nachtsheim, M., Meinecke, P., Morlot, S., et al. (2007). Expansion of the genotypic and phenotypic spectrum in patients with KRAS germline mutations. *J Med Genet* 44, 131-135.
460. Bertola, D.R., Pereira, A.C., Brasil, A.S., Albano, L.M., Kim, C.A., and Krieger, J.E. (2007). Further evidence of genetic heterogeneity in Costello syndrome: involvement of the KRAS gene. *J Hum Genet* 52, 521-526.
461. Stark, Z., Gillessen-Kaesbach, G., Ryan, M.M., Cirstea, I.C., Gremer, L., Ahmadian, M.R., Savarirayan, R., and Zenker, M. (2012). Two novel germline KRAS mutations: expanding the molecular and clinical phenotype. *Clin Genet* 81, 590-594.
462. Ekval, S., Wilbe, M., Dahlgren, J., Legius, E., van Haeringen, A., Westphal, O., Anneren, G., and Bondeson, M.L. (2015). Mutation in NRAS in familial Noonan syndrome--case report and review of the literature. *BMC Med Genet* 16, 95.
463. Kerr, B., Delrue, M.A., Sigaudy, S., Perveen, R., Marche, M., Burgelin, I., Stef, M., Tang, B., Eden, O.B., O'Sullivan, J., et al. (2006). Genotype-phenotype correlation in Costello syndrome: HRAS mutation analysis in 43 cases. *J Med Genet* 43, 401-405.
464. Zampino, G., Pantaleoni, F., Carta, C., Cobellis, G., Vasta, I., Neri, C., Pogna, E.A., De Feo, E., Delogu, A., Sarkozy, A., et al. (2007). Diversity, parental germline origin, and phenotypic spectrum of de novo HRAS missense changes in Costello syndrome. *Hum Mutat* 28, 265-272.
465. Gripp, K.W., Hopkins, E., Sol-Church, K., Stabley, D.L., Axelrad, M.E., Doyle, D., Dobyns, W.B., Hudson, C., Johnson, J., Tenconi, R., et al. (2011). Phenotypic analysis of individuals with Costello syndrome due to HRAS p.G13C. *Am J Med Genet A* 155A, 706-716.
466. Martinelli, S., Krumbach, O.H.F., Pantaleoni, F., Coppola, S., Amin, E., Pannone, L., Nouri, K., Farina, L., Dvorsky, R., Lepri, F., et al. (2018). Functional Dysregulation of CDC42 Causes Diverse Developmental Phenotypes. *Am J Hum Genet*.
467. Wiel, L., Venselaar, H., Veltman, J.A., Vriend, G., and Gilissen, C. (2017). Aggregation of population-based genetic variation over protein domain homologues and its potential use in genetic diagnostics. *Hum Mutat* 38, 1454-1463.
468. Lelieveld, S.H., Wiel, L., Venselaar, H., Pfundt, R., Vriend, G., Veltman, J.A., Brunner, H.G., Vissers, L., and Gilissen, C. (2017). Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am J Hum Genet* 101, 478-484.
469. Geisheker, M.R., Heymann, G., Wang, T., Coe, B.P., Turner, T.N., Stessman, H.A.F., Hoekzema, K., Kvarnung, M., Shaw, M., Friend, K., et al. (2017). Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat Neurosci* 20, 1043-1051.
470. Colby, W.W., Hayflick, J.S., Clark, S.G., and Levinson, A.D. (1986). Biochemical characterization of polypeptides encoded by mutated human Ha-ras1 genes. *Mol Cell Biol* 6, 730-734.
471. Goriely, A., and Wilkie, A.O. (2010). Missing heritability: paternal age effect mutations and selfish spermatogonia.

- Nat Rev Genet 11, 589.
472. Goriely, A., Hansen, R.M., Taylor, I.B., Olesen, I.A., Jacobsen, G.K., McGowan, S.J., Pfeifer, S.P., McVean, G.A., Rajpert-De Meyts, E., and Wilkie, A.O. (2009). Activating mutations in FGFR3 and HRAS reveal a shared genetic origin for congenital disorders and testicular tumors. *Nat Genet* 41, 1247-1252.
 473. Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat Rev Genet* 11, 415-425.
 474. Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 12, 745-755.
 475. Jansen, S., Hoischen, A., Coe, B.P., Carvill, G.L., Van Esch, H., Bosch, D.G.M., Andersen, U.A., Baker, C., Bauters, M., Bernier, R.A., et al. (2017). A genotype-first approach identifies an intellectual disability-overweight syndrome caused by PHIP haploinsufficiency. *Eur J Hum Genet*.
 476. Fergelot, P., Van Belzen, M., Van Gils, J., Afenjar, A., Armour, C.M., Arveiler, B., Beets, L., Burglen, L., Busa, T., Collet, M., et al. (2016). Phenotype and genotype in 52 patients with Rubinstein-Taybi syndrome caused by EP300 mutations. *Am J Med Genet A* 170, 3069-3082.
 477. Field, M., Tarpey, P.S., Smith, R., Edkins, S., O'Meara, S., Stevens, C., Tofts, C., Teague, J., Butler, A., Dicks, E., et al. (2007). Mutations in the BRWD3 gene cause X-linked mental retardation associated with macrocephaly. *Am J Hum Genet* 81, 367-374.
 478. Barsh, G.S., Farooqi, I.S., and O'Rahilly, S. (2000). Genetics of body-weight regulation. *Nat New Biol* 404, 644-651.
 479. Stunkard, A.J., Harris, J.R., Pedersen, N.L., and McClearn, G.E. (1990). The body-mass index of twins who have been reared apart. *N Engl J Med* 322, 1483-1487.
 480. Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nat New Biol* 518, 197-206.
 481. Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., Bradfield, J.P., Esko, T., Giri, A., Graff, M., et al. (2018). Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. *Nat Genet* 50, 26-41.
 482. Network, E.R. <https://endo-ern.eu/ern/>.
 483. Sifrim, A., Hitz, M.P., Wilsdon, A., Breckpot, J., Turki, S.H., Thienpont, B., McRae, J., Fitzgerald, T.W., Singh, T., Swaminathan, G.J., et al. (2016). Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat Genet* 48, 1060-1065.
 484. Cram, D.S., and Zhou, D. (2016). Next generation sequencing: Coping with rare genetic diseases in China. *Intractable Rare Dis Res* 5, 140-144.
 485. K Vervier, J.M. (2017). TiSAn: Tissue Specific Variant Annotation. *BioRxiv*.
 486. Quinodoz, M., Royer-Bertrand, B., Cisarova, K., Di Gioia, S.A., Superti-Furga, A., and Rivolta, C. (2017). DOMINO: Using Machine Learning to Predict Genes Associated with Dominant Disorders. *Am J Hum Genet* 101, 623-629.
 487. Zhang, X., Li, M., Lin, H., Rao, X., Feng, W., Yang, Y., Mort, M., Cooper, D.N., Wang, Y., Wang, Y., et al. (2017). regSNPs-splicing: a tool for prioritizing synonymous single-nucleotide substitution. *Hum Genet* 136, 1279-1289.
 488. Global Alliance for, G., and Health. (2016). GENOMICS. A federated ecosystem for sharing genomic, clinical data.

- Science (80-) 352, 1278-1280.
489. Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23, 185-188.
 490. Jansen, S., Kleefstra, T., Willemsen, M.H., de Vries, P., Pfundt, R., Hehir-Kwa, J.Y., Gilissen, C., Veltman, J.A., de Vries, B.B., and Vissers, L.E. (2016). De novo loss-of-function mutations in X-linked SMC1A cause severe ID and therapy-resistant epilepsy in females: expanding the phenotypic spectrum. *Clin Genet* 90, 413-419.
 491. Moortgat, S., Berland, S., Aukrust, I., Maystadt, I., Baker, L., Benoit, V., Caro-Llopis, A., Cooper, N.S., Debray, F.G., Faivre, L., et al. (2017). HUWE1 variants cause dominant X-linked intellectual disability: a clinical study of 21 patients. *Eur J Hum Genet*.
 492. de Lange, I.M., Helbig, K.L., Weckhuysen, S., Moller, R.S., Velinov, M., Dolzhanskaya, N., Marsh, E., Helbig, I., Devinsky, O., Tang, S., et al. (2016). De novo mutations of KIAA2022 in females cause intellectual disability and intractable epilepsy. *J Med Genet* 53, 850-858.
 493. Cotton, A.M., Chen, C.Y., Lam, L.L., Wasserman, W.W., Kobor, M.S., and Brown, C.J. (2014). Spread of X-chromosome inactivation into autosomal sequences: role for DNA elements, chromatin features and chromosomal domains. *Hum Mol Genet* 23, 1211-1223.
 494. Marks, H., Kerstens, H.H., Barakat, T.S., Splinter, E., Dirks, R.A., van Mierlo, G., Joshi, O., Wang, S.Y., Babak, T., Albers, C.A., et al. (2015). Dynamics of gene silencing during X inactivation using allele-specific RNA-seq. *Genome Biol* 16, 149.
 495. Ewans, L.J., Field, M., Zhu, Y., Turner, G., Leffler, M., Dinger, M.E., Cowley, M.J., Buckley, M.F., Scheffer, I.E., Jackson, M.R., et al. (2017). Gonadal mosaicism of a novel IQSEC2 variant causing female limited intellectual disability and epilepsy. *Eur J Hum Genet*.
 496. Fieremans, N., Van Esch, H., de Ravel, T., Van Driessche, J., Belet, S., Bauters, M., and Froyen, G. (2015). Microdeletion of the escape genes KDM5C and IQSEC2 in a girl with severe intellectual disability and autistic features. *Eur J Med Genet* 58, 324-327.
 497. Helm, B.M., Powis, Z., Prada, C.E., Casasbuenas-Alarcon, O.L., Balmakund, T., Schaefer, G.B., Kahler, S.G., Kaylor, J., Winter, S., Zarate, Y.A., et al. (2017). The role of IQSEC2 in syndromic intellectual disability: Narrowing the diagnostic odyssey. *Am J Med Genet A* 173, 2814-2820.
 498. Tzschach, A., Grasshoff, U., Beck-Woedl, S., Dufke, C., Bauer, C., Kehrer, M., Evers, C., Moog, U., Oehl-Jaschowitz, B., Di Donato, N., et al. (2015). Next-generation sequencing in X-linked intellectual disability. *Eur J Hum Genet* 23, 1513-1518.
 499. Miller, D.T., Adam, M.P., Aradhya, S., Biesecker, L.G., Brothman, A.R., Carter, N.P., Church, D.M., Crolla, J.A., Eichler, E.E., Epstein, C.J., et al. (2010). Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86, 749-764.
 500. Bi, W., Borgan, C., Pursley, A.N., Hixson, P., Shaw, C.A., Bacino, C.A., Lalani, S.R., Patel, A., Stankiewicz, P., Lupski, J.R., et al. (2013). Comparison of chromosome analysis and chromosomal microarray analysis: what is the value of chromosome analysis in today's genomic array era? *Genet Med* 15, 450-457.
 501. Coe, B.P., Witherspoon, K., Rosenfeld, J.A., van Bon, B.W., Vulto-van Silfhout, A.T., Bosco, P., Friend, K.L., Baker, C., Buono, S., Vissers, L.E., et al. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 46, 1063-1071.

502. Gambin, T., Yuan, B., Bi, W., Liu, P., Rosenfeld, J.A., Coban-Akdemir, Z., Pursley, A.N., Nagamani, S.C.S., Marom, R., Golla, S., et al. (2017). Identification of novel candidate disease genes from de novo exonic copy number variants. *Genome Med* 9, 83.
503. Reyes-Palomares, A., Bueno, A., Rodriguez-Lopez, R., Medina, M.A., Sanchez-Jimenez, F., Corpas, M., and Ranea, J.A. (2016). Systematic identification of phenotypically enriched loci using a patient network of genomic disorders. *BMC Genomics* 17, 232.
504. Conte, F., Oti, M., Dixon, J., Carels, C.E., Rubini, M., and Zhou, H. (2016). Systematic analysis of copy number variants of a large cohort of orofacial cleft patients identifies candidate genes for orofacial clefts. *Hum Genet* 135, 41-59.
505. Tsuchida, N., Nakashima, M., Kato, M., Heyman, E., Inui, T., Haginoya, K., Watanabe, S., Chiyonobu, T., Morimoto, M., Ohta, M., et al. (2017). Detection of copy number variations in epilepsy using exome data. *Clin Genet*.
506. Sims, D., Sudbery, I., Illott, N.E., Heger, A., and Ponting, C.P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet* 15, 121-132.
507. Kalari, K.R., Casavant, M., Bair, T.B., Keen, H.L., Comerón, J.M., Casavant, T.L., and Scheetz, T.E. (2006). First exons and introns--a survey of GC content and gene structure in the human genome. *In Silico Biol* 6, 237-242.
508. Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17, 333-351.

Nederlandse samenvatting

Neurobiologische ontwikkelingsstoornissen of ‘Neurodevelopmental disorders (NDDs)’ zijn een groep van stoornissen die gekenmerkt worden door een verstoorde ontwikkeling op jonge leeftijd. Daardoor ontstaan problemen met persoonlijk, sociaal, academisch of beroepsmatig functioneren. Aandoeningen binnen de groep NDDs zijn onder andere verstandelijke beperking en een stoornis in het autistisch spectrum. Het is bekend dat genetische afwijkingen met regelmaat bijdragen aan het ontstaan van NDDs, met name als de ontwikkelingsstoornissen gepaard gaan met andere symptomen zoals hersenafwijkingen, epilepsie, aangeboren afwijkingen of bijzondere gelaatskenmerken. Door een onderliggende genetische aandoening te diagnosticeren kan meer inzicht worden verkregen in de prognose en het te volgen behandeltraject. Daarnaast verschaft het duidelijkheid over het herhalingsrisico voor ouders op het krijgen van een tweede kind met de aandoening. Tot slot moet de emotionele impact van een genetische diagnose niet worden onderschat: veel ouders van kinderen met ontwikkelingsstoornissen zijn al jarenlang op zoek naar een oorzaak van de problemen en kunnen dankzij een diagnose de situatie beter accepteren.

Genetisch testen stamt uit de jaren '70. Echter, met de beperkte technische mogelijkheden bleven veel patiënten met NDD nog lange tijd zonder diagnose. Nieuwe technologische ontwikkelingen hebben hierin verandering gebracht. Na de introductie van array-technieken werd in 2010 Next Generation Sequencing (NGS) voor het eerst succesvol gebruikt om genen op te sporen die betrokken zijn bij een Mendeliaans overervende ziekte. Anno 2018 zijn NGS-technieken, en dan met name Whole Exome Sequencing (WES), steeds beter geworden. Grote aantallen nieuwe genen - en daarmee de genetische diagnoses - zijn hiermee geïdentificeerd. Dit geldt ook voor genen geassocieerd met NDDs, welke in dit proefschrift ‘next generation NDDs’ worden genoemd. Verandering in deze genen, mutaties genoemd, gaan vaak gepaard met het voorkomen van verschillende andere symptomen. Al deze symptomen samen vormen een syndroom.

Toch kent ook NGS zijn beperkingen. Met NGS kan nog steeds niet bij alle patiënten een diagnose worden gesteld: bij ~50% van de NDD-patiënten bij wie NGS wordt toegepast, wordt een (mogelijke) genetische oorzaak gevonden. Daarnaast is er voor patiënten gediagnosticeerd met een next generation NDD beperkte informatie beschikbaar over klinische, prognostische en eventuele behandel adviezen. Tot slot worden soms mutaties gevonden in genen waarover vooralsnog weinig bekend is. Het blijft dan onduidelijk of de mutaties daadwerkelijk pathogeen (ziekmakend) zijn. Deze beperkingen zijn voor artsen en hun patiënten en hun ouders lastig. Daarom is dit proefschrift erop gericht om meer inzicht te verkrijgen in de genetische, klinische en biologische aspecten van next-generation NDDs. Om dit te verwezenlijken is een groot cohort (onderzoeksgroep) van 826 patiënten

met onbegrepen NDDs samengesteld. Kenmerkend aan de patiënten is dat zij allemaal een verstandelijke beperking hebben. Daarnaast komt een verscheidenheid aan andere klinische kenmerken voor, zoals aangeboren afwijkingen, epilepsie of een afwijkende groei.

Genetische aspecten

Bij alle patiënten in het cohort werd NGS gedaan. Dankzij de grote hoeveelheid patiënten bij wie dit werd toegepast en door gebruik te maken van alternatieve manieren om de NGS-data te interpreteren, konden mutaties in nieuwe genen en daarbij behorende syndromen worden ontdekt. Zo werden in hoofdstuk 3 met behulp van een statistische benadering en vergelijking met de literatuur tien nieuwe kandidaat genen gevonden. In een ander hoofdstuk, hoofdstuk 7, werd een andere benadering gebruikt: door te kijken naar een honderdtal genen die coderen voor eiwitten die nauw met elkaar samenwerken binnen een zogenaamde 'pathway', werd de aandacht gevestigd op een gen dat voorheen nooit werd geassocieerd met NDDs: het *RHEB*-gen. Patiënten met een mutatie in dit gen hadden allemaal een abnormaal groot hoofd (macrocefalie), een kenmerk dat reeds sterk geassocieerd was met de onderzochte pathway.

Klinische aspecten

Door de klinische (medische) kenmerken van de patiënten – de fenotypes – met elkaar te vergelijken, is in dit proefschrift verder bewijs gevonden dat de ontdekte kandidaat genen daadwerkelijk kunnen worden gerekend tot de next generation NDDs. Daarnaast is zowel nationaal als internationaal gezocht naar extra patiënten. Meer patiënten leverde niet alleen meer bewijs voor de NDD-kandidaat genen, maar zorgde er ook voor dat meer inzicht verworven kon worden in het klinisch beeld passend bij de nieuwe syndromen. Zo werd in hoofdstuk 9 duidelijk dat mutaties in het *USP9X*-gen bij meisjes geassocieerd zijn met niet alleen een ontwikkelingsstoornis, maar ook met een verscheidenheid aan herkenbare aangeboren afwijkingen. Het *TLK2*-gen in hoofdstuk 5 daarentegen wordt in verband gebracht met een minder specifiek en minder herkenbaar fenotype. Echter, doordat een relatief grote hoeveelheid patiënten (40) in de wereld werd ontdekt, was er voldoende bewijs om ook het syndroom geassocieerd met dit gen aan te merken als next generation NDD.

Biologische aspecten

Voor sommige kandidaat genen bleef de hoeveelheid patiënten, ondanks het wereldwijd delen van data, beperkt: slechts drie patiënten hadden mutaties in het *RHEB*-gen (hoofdstuk 7) en vijf hadden mutaties in het *RAB11B*-gen (hoofdstuk 8). Verder bewijs voor deze genen als next generation NDDs werd niet alleen geleverd door de fenotypes van de patiënten met elkaar te vergelijken, maar ook door aan te tonen dat de mutaties leidden tot verstoring van (cel)biologische processen. Ook voor andere nieuwe genen werden aanvullende functionele studies in dier- en cel-modellen gedaan. Dit leverde nieuwe inzichten in (verstoorde) onderliggende cel-processen op. Zo hadden zebrafissen met specifieke mutaties in het

RAC1-gen een verminderde deling van cellen in de hersenen (hoofdstuk 6) en liet onderzoek in cellen van patiënten met mutaties in het *SON*-gen zien dat niet alleen het *SON*-eiwit, maar ook andere belangrijke eiwitten verstoord waren (hoofdstuk 4). In hoofdstuk 7 werd binnen het set aan genen die zijn gerelateerd aan het *MTOR*-gen, een relatie gelegd tussen de hoofdomtrek van de algemene bevolking en de opvallende hoeveelheid patiënten met een macrocefalie.

Neurodevelopmental disorder: a next generation – what's next?

In dit proefschrift konden uiteindelijk 14 genen worden toegevoegd aan de NDD-gen lijst: zeven genen als kandidaat genen (hoofdstuk 3) en 7 genen met bijbehorende syndromen als nieuwe next generation NDDs (hoofdstuk 4-10). Deze nieuwe syndromen hebben, net als andere next generation NDDs tenminste één kenmerk gemeen: ze zijn allemaal zeer zeldzaam. De beschreven aantallen patiënten blijven daarom, ondanks uitgebreide (inter) nationale samenwerkingen, beperkt. Naast het verder verbeteren van mogelijkheden om data te delen, zullen toekomstige klinische vervolgstudies, zoals in hoofdstuk 11 van dit proefschrift voor PURA syndroom is gedaan, van groot belang zijn. Hiermee kan meer inzicht worden verworven in het klinisch spectrum dat past bij de nieuwe syndromen. Tevens zullen steeds grotere cohorten van patiënten nodig zijn om ook de nóg zeldzamere syndromen te kunnen opsporen.

Een tweede opvallende overeenkomst tussen de verschillende syndromen beschreven in dit proefschrift, is de variabiliteit in ernst van symptomen binnen de syndromen. Bij de selectie van het cohort voor dit proefschrift is er bewust voor gekozen niet alleen de ernstiger aangedane patiënten te includeren, maar ook patiënten met een milder fenotype. Door ook deze patiënten te selecteren werd in dit proefschrift niet alleen aangetoond dat de NGS-opbrengst voor deze patiënten vergelijkbaar was met ernstiger aangedane patiënten, maar ook dat mutaties in bepaalde genen bij sommige patiënten kunnen resulteren in zeer ernstige ontwikkelingsstoornissen en bij anderen tot veel mildere problemen. Juist voor de patiënten met een milder fenotype moet in de toekomst rekening worden gehouden met een overervingspatroon dat anders is dan bij patiënten met een ernstige ontwikkelingsstoornis. Mutaties zullen niet altijd meer ‘*de novo*’ zijn (nieuw ontstaan in de patiënt, niet aanwezig in beide ouders), maar zullen soms ook worden teruggevonden bij een vader of moeder die, net als hun kind, milde problemen heeft.

Functionele experimenten zijn van essentieel belang om de zeldzame nieuwe next-generation NDDs te bevestigen en onderliggende ziektemechanismen te begrijpen. Uiteindelijk zal deze kennis nodig zijn om het ultieme doel te behalen: de ontwikkeling van therapeutische middelen. Veel werk werd door collegae verzet om te komen tot de resultaten in dit proefschrift getoond. Om dit in de toekomst voor alle next-generation NDDs tot in detail te kunnen doen, zal veel geld en tijd nodig zijn: zaken waar helaas grote tekorten in zijn.

Opnieuw zullen goede (inter)nationale samenwerkingen essentieel zijn. In landen zoals de VS en Canada, maar meer recenter ook in de EU, zijn middels samenwerkingsverbanden hiervoor de eerste stappen gezet.

Aankomende jaren zal de zoektocht naar nieuwe next-generation NDDs verder gaan. Het verbeteren van WES en het analyseren van data op alternatieve manieren zal waarschijnlijk leiden tot de ontdekking van meer next-generation NDDs. Daarnaast zal de implementatie van Whole Genome Sequencing (WGS) in de diagnostiek tekortkomingen van WES kunnen wegnemen. Echter, de focus zal niet alleen moeten liggen in het vinden van diagnoses voor patiënten die vooralsnog geen diagnose hebben. Het zal een hele klus zijn om goede patiëntenzorg op te zetten voor patiënten gediagnosticeerd met next-generation NDDs. De rol van ouders in dit proces zal daarbij zeer belangrijk zijn.

Dankwoord

Bijna vijf jaar geleden leek dit moment nog zo ver weg: het schrijven van een dankwoord voor mijn proefschrift. Een perfect moment om te reflecteren op de afgelopen tijd en mezelf te realiseren hoeveel mensen hebben bijgedragen aan niet alleen dit proefschrift, maar ook aan de fantastische onderzoekstijd die ik heb gehad.

Promoveren – de (co)promotoren

Januari 2014 begon ik, net terug van mijn laatste coschap in Indonesië, aan een nieuwe uitdaging: promotieonderzoek op de afdeling klinische genetica. Stiekem moet ik bekennen dat ik eigenlijk niet eens wist wat zo'n promotie nu precies inhield. Ik vond onderzoek doen én de genetica gewoon erg leuk. Het leek me dus de ideale baan. **Han**, vanaf het begin af aan ben je betrokken geweest als mijn promotor. Ik kan me de eerste keer nog herinneren dat ik een overleg gepland had staan met je. Ik was goed voorbereid en had urenlang gewerkt aan het schrijven van mijn 'Training and Supervision Plan'. Dit moest natuurlijk door jou als promotor goedgekeurd en ondertekend worden. We spraken er eens goed over, waarna je vervolgens het TSP ondertekende zonder ook maar één woord te lezen van wat erin stond. Met natuurlijk wel de belangrijke boodschap aan mij dat ik zelf nooit iets zou moeten ondertekenen zonder het zelf gelezen te hebben. En dat je er vertrouwen in had dat het TSP overeenkwam met wat we net besproken hadden. Terugkijkend was dit eigenlijk een belangrijk moment: je sprak je vertrouwen uit in mij. Ik ben je erg dankbaar dat je dit vertrouwen altijd hebt gehouden en hebt uitgesproken op momenten dat het nodig was. Je hebt me de vrijheid gegeven om mijn eigen weg te zoeken als onderzoeker, en stuurde me de goede richting uit en inspireerde me op het juiste moment. Na verloop van tijd begon ik me te realiseren dat de stiltes die vielen tijdens een overleg helemaal niet ongemakkelijk waren: meestal kwam je na zo'n stilte met een perfecte opmerking waardoor ik weer weken verder kon. Dank voor niet alleen de wetenschappelijk input, maar ook de persoonlijke interesse die je altijd hebt getoond. **Tjitske**, aan het einde van mijn eerste jaar werd jij mijn copromotor en sloot jij aan bij de overleggen tussen Han en mij. Wat hebben we ons vaak samen afgevraagd of Han toch nog wel kwam opdagen voor ons overleg. Maar meestal bleek hij niet zo ver weg te zijn en stond hij een paar deuren verder nog snel even iets te overleggen. We hadden samen ook onze wekelijkse gesprekken. Het was fijn om met jou te brainstormen over de grote lijnen van mijn boekje en de mogelijkheden voor nieuwe onderzoeksprojecten. De combinatie van onderzoek en patiëntenzorg heeft voor mij het promotieonderzoek extra glans gegeven. Persoonlijk vond ik tijdens mijn onderzoeksperiode niets mooier dan ouders te vertellen dat we na jaren zoeken eindelijk een diagnose gevonden hadden. Jij hebt me laten zien dat patiëntenzorg niet ten koste hoeft te gaan van onderzoek: je loopt voorop in onderzoek, maar met hart voor de patiënten. Dit heb ik – en doe ik nog steeds – altijd ontzettend bewonderd. Ik ben ervan overtuigd dat je dit aankomende jaren met veel succes

gaat voortzetten! Net als Han was jij, **Lisenka**, vanaf het begin bij mijn promotietraject betrokken. Het WAC-artikel was het eerste project waar ik samen met jou aan werkte. Ik kan me nog herinneren dat het moment was gekomen om het artikel te submitten. Mijn eerste artikel! Ik had geen idee hoe het submitten van een artikel in zijn werk ging, dus deden we het samen. De 'druk op de knop' werd uiteraard vastgelegd. Dit is één van de voorbeelden die ik zo gewaardeerd heb aan je de afgelopen jaren: we zagen en spraken elkaar niet dagelijks, maar als ik een luisterend oor of ergens hulp bij nodig had kon ik altijd bij je terecht. Daarnaast waren je kritische blik en je oog voor detail heel fijn. Niet alleen tijdens het uitvoeren van het onderzoek, maar ook tijdens het schrijven van artikelen en uiteindelijk mijn proefschrift hield je me scherp met je nuttige commentaren. Heel veel dank hiervoor!

Samenwerking met verschillende professies

Al snel werd mij duidelijk dat ik als PhD student was begonnen op een hele grote afdeling, waar mensen werkten met compleet verschillende achtergronden. In het begin was dit helemaal niet zo gemakkelijk: zo heb ik als één van de enige arts-onderzoekers toch menigmaal uitgelegd aan collega-PhD's waar de afdeling klinische genetica zich precies bevond en wat wij daar precies deden. Daarnaast moet ik eerlijk toegeven dat ik weinig snapte van wat er verder 'in het lab' gebeurde. Ik realiseerde me dat ik zóveel te leren had. Tijdens mijn onderzoekstijd heb ik de mogelijkheid gehad me te verdiepen en verder te denken dan waar ik voor opgeleid was. Dank aan alle **collega's in de research en diagnostiek** voor het altijd geven van uitleg, nuttige feedback en het helpen oplossen van moeilijk vragen. Dankzij jullie is mijn kijk op de genetica enorm verbreed. **Ronald**, we hebben afgelopen jaren samen aan heel wat projecten gewerkt. Ons eerste project, USP9X, was meteen raak. Op het moment dat ik je vertelde dat we USP9X-patiënten hadden met een ciliopathie-achtig fenotype was je meteen een en al oor. Of ik wist dat USP9X hoog stond op je ciliopathie-kandidatenlijst? Het werd een fantastisch project waardoor ik niet alleen ontzettend veel heb kunnen leren over functionele studies en ciliopathieën, maar ook je onderzoeksgroep heb leren kennen. Ook al was ik nooit officieel onderdeel van je onderzoeksgroep, ik heb me altijd welkom gevoeld! Dank voor de interesse die je altijd hebt getoond en de hele fijne samenwerking die we altijd hebben gehad. Hopelijk kunnen we hier in de toekomst een vervolg aan geven! **Brooke**, it was a great pleasure to work together with you on the USP9X project. We struggled about strange lab findings, discussed difficult results, and thought about next steps in our project: by combining our results we finally published a paper I'm very proud of! **Ideke**, je was mijn maatje in het RAB11B project. Hoe vaak hebben we samen rondgelopen op zoek naar Ronald? Er lag tijdens het project een behoorlijk druk op het afronden van je proefschrift. Maar voor ik het wist was het af, inclusief supermooie data voor ons artikel! **Marieke, Janita en Barbara**, jullie hebben mij kennis laten maken met de wereld van GWAS. Dankzij jullie uitleg en hulp ging ik het uiteindelijk echt begrijpen. **Rolph**, jij was gedurende mijn gehele PhD tijd mijn vraagbaak in de diagnostiek. Ik ben altijd onder de indruk geweest van de hoeveelheid werk jij verzet.

Ik kon altijd bij je binnenlopen en je nam, hoe druk je het ook had, altijd de tijd om mijn vraag te beantwoorden. Kun je het je nog herinneren dat ik aan je vroeg waar ik missende dossiers kon vinden? Uiteindelijk zijn we samen de hele afdeling over gegaan op zoek naar missende dossiers om consenten te achterhalen. Dank voor de energie die je altijd in de verschillende projecten hebt gestopt! **Nicole**, jij was degene die me meenam naar Goldrain, waar ik een ontzettend leuke week gehad heb. Dank je wel voor oprechte interesse en je altijd scherpe blik op mijn onderzoeksresultaten. **Prof. Katsanis**, thank you for several fruitful collaborations! **Dorien, Annette and Misa**: thanks for the great work on the WAC-gene! **Christian en Stefan**, jullie fantastische werk droeg bij aan de basis van mijn proefschrift. Veel dank! **Prof. Wilkie**, many thanks for the great collaboration and your always valuable comments on the TLK2-project. **Maria**, it was nice to have drinks at the ASHG and to see our projects progressing and improving every year. Thanks for a fantastic collaboration! **Lachlan**, it was great to work with you on USP9X. Your extensive knowledge about this gene was very helpful. **Geeske en Ype**, het was ontzettend leuk om met jullie samen te werken. Jullie input voor ons MTOR-artikel en kennis van de pathway is ontzettend waardevol geweest. We gaan elkaar vast nog regelmatig tegenkomen! **Dr. Mancini**, bedankt voor je hulp bij het beoordelen van MRI-beelden en de heldere uitleg die je mij gaf. **Dr. Banka**, it was a pleasure to work together with you and your colleagues on the RAC1-project, thank you for the successful collaboration. **Dr. Nellaker**, thanks for your interesting work on the PURA- and TLK-faces. **Jung-Hyun, Deepali and Erin**: it was a special moment to meet you at the AHSG conference in Vancouver, after so many conversations by email. Thanks for your hard work on the SON gene!

Bijzonder contact met patiënten en hun ouders

Een van het meest inspirerende en motiverende ervaringen tijdens mijn PhD traject was het werken met **patiënten en hun ouders**. Ik vond het bijzonder om te zien dat ouders zo open stonden om bij te dragen aan de wetenschap. Mede dankzij hen is dit proefschrift geworden zoals het nu is. Hopelijk kunnen veel andere artsen, ouders en hun kinderen gebruik maken van de kennis die dankzij patiënten en hun ouders uiteindelijk op papier is gezet. **Mel and Ceciel**, you are definitely the most special parents I've met during my PhD project. It is unbelievable what you have done (and are still doing!) for all patients with PURA syndrome, including your daughters. Mel, I'll never forget the moment we met for the first time in the car on our way to the first PURA conference. I had no idea what to expect and I was overwhelmed by your enthusiasm. I'm so glad I decided to stay in the car! Ceciel, it was exactly that day we met for the first time as well: waiting for hours on the airport and in the airplane (without any food...). We had so much fun when we heard your name on the speaker for a last call. All other members of the **PURA Foundation board**: thanks for keeping a fantastic organization going and for your interest in my work and me as person. Many thanks as well for the **PURA Syndrome Global Research Network**: I feel blessed to be involved in such an inspiring network of researchers and clinicians. It's always a pleasure meeting you at

the PURA conferences and to discuss future directions of the research team. **Dierk**, thanks for inviting me in Munich, it was great to see your lab and to meet your team. **David and Diana**, thanks to you I became involved in PURA Syndrome. I'm very proud on the results of our work we published last year. It was really a pleasure to work with you. I'm convinced many more collaborations will follow. **Jennifer**, thanks for the inspiring conversations with drinks at the bar! Belangrijk voor mijn werk in de kliniek was de **afdeling klinische genetica**. Dank aan alle **artsen** voor het geven van de mogelijkheid om aan te sluiten bij polibezoeken van patiënten en het helpen om medische gegevens compleet te krijgen. Dames van het **secretariaat**, het was ontzettend fijn dat ik altijd bij jullie terecht kon als er iets geregeld moest worden, geen vraag was te gek voor jullie! Tot slot wil ik alle **collega's van andere ziekenhuizen** bedanken voor het zorgvuldig verzamelen van klinische details en het meewerken aan mijn onderzoek.

Congressen en cursussen... een scala aan anekdotes

Een hele leuke bijkomstigheid van het promotietraject dat ik doorlopen heb, zijn wat mij betreft de vele cursussen en congressen in binnen- en buitenland die ik heb mogen bezoeken. Het was niet alleen de perfecte manier om nieuwe kennis op te doen en buitenlandse collega's te ontmoeten, maar ook een bijzondere manier om mijn Nijmeegse collega's beter te leren kennen. Wat een hoop goede herinneringen heb ik hieraan...! Ik kijk met veel plezier terug op de congressen waar ik met collega's verbleef in de 'Nijmegen-villa'. Dank aan **alle collega's** met wie ik met veel plezier samen in één van deze huizen heb gezeten. **Alex**, het mooiste huisfeest dat jij organiseerde en waar ik bij ben geweest is toch wel het feest in Glasgow geweest. Tot diep in de nacht hebben we in 'onze' villa gefeest (en **Rolph**, jouw binnenkomst was wat mij betreft memorabel: trots showde je de verschillende flessen drank in de binnenzakken van je jas). **Sandra**, in Barcelona deelde ik een appartement onder andere met jou. Het was ontzettend gezellig, we hebben uiteindelijk over de meest onzinnige dingen zitten kletsen. **Anneke**, we waren samen naar de Manchester Dysmorphology course. We liepen 's ochtends samen naar de cursusruimte omdat we geen zin hadden om met de bus te gaan. Ik vergeet nooit dat je toen opmerkte dat je je verbaasde over het feit dat ik net zo snel liep als jij. In het ziekenhuis liep ik namelijk altijd nogal traag. Erg grappig vond ik dat en typisch een opmerking voor jou. **Nicole**, aan het einde van een late sessie op de ASHG in Vancouver bleken alle andere collega's al lang en breed te hebben gegeten. Bij een goede Italiaan hebben we uiteindelijk een hele gezellige avond gehad. **Han**, tijdens het congres in Kopenhagen zou jij nog even naar mijn presentatie kijken. Dus stonden we daar, midden in de congreshal, mijn presentatie door te nemen terwijl we zo ongeveer iedere minuut werden onderbroken doordat iemand een praatje met je kwam maken. Maar ondertussen heb je me toen wel een aantal presentatie-skills bijgebracht waar ik nu nog altijd gebruik van maak. **Ilja**, wij waren samen in Cardiff. Het was toch hilarisch dat wij iedere ochtend een half uur eerder gingen ontbijten omdat we ons tegenover onze andere collega's schaamden dat wij zoveel aten van het enorme ontbijtbuffet. Tegen de

tijd dat de andere collega's arriveerden waren wij net begonnen aan onze derde ronde 😊.

Stefan, weet je nog dat we samen met Lot super zenuwachtig die enorme congreshal op de ESHG binnenstapten voor ons eerste praatje? Wat een ontlading was het toen we alle drie ons praatje achter de rug hadden. We hebben in heel wat Nijmegen-huizen samen gezeten. Orlando was toch wel een favoriet... 'Eitje met bacon?'. Na het congres in Glasgow hebben we nog een paar dagen heel veel lol gehad in Edinburgh. Een laatste anekdote, die totaal niet in deze alinea past maar me nu toch te binnenschiet: de enorme chocoladetaart die je aan me gaf toen ik al die consenten had nagekeken. We hebben meerdere koffiepauzes op die dag moeten inlassen om hem op te krijgen! **Laurens en Jacob**, volgens mij was Kopenhagen ons eerste congres samen. Ik was niet helemaal in goede doen, mijn arm was net uit het gips. Jullie was zo ontzettend behulpzaam! Laurens, een half jaar later zijn we ook samen naar Orlando geweest. Aansluitend aan het congres zijn we naar Universal Studios geweest. De afterparty met Harry Potter films moeten we nu toch echt maar eens organiseren!

Christian, jij was altijd zo ontzettend hard aan het werk tijdens congressen, ongelofelijk hoe je dit altijd volhoudt. Want naast het werk mis je nooit eens een borrel of feestje. Na het congres in Baltimore hebben wij samen met Lot een bus gepakt voor een weekje New York (waar we overigens het appartement van **Susanne** hebben mogen bewonderen). Wat ontzettend gaaf was dit! En heel erg leuk dat Michiel en jij elkaar daar hebben leren kennen.

Lot, onze eerste buitenland-ervaring was Bertinoro. Wat een goede herinneringen heb ik hieraan! Want ondanks dat we beiden op dezelfde afdeling werkten, leerden we elkaar hier pas echt kennen. Vanaf dat moment zijn we zo vaak samen (natuurlijk tot in de puntjes en alles op tijd voorbereid...) naar congressen geweest. Jij zei altijd dat je het zo fijn vond als alles 'lekker ongecompliceerd' was. Nou, volgens mij is dat wel gelukt!

Fantastische collega's

Als je meer dan vier jaar rondloopt op de werkvloer leer je natuurlijk veel directe collega's steeds beter kennen. Tot mijn plezier werd de groep onderzoekers binnen de Klinische Genetica steeds groter. **Aisha, Anneke, Arjan, Chantal, Daniëlle, Gijsbert, Janet, Illja, Ingrid, Linde, Lot, Marc, Sandra en Yvonne**, dank jullie wel voor jullie support en gezelschap. **Aisha**, het was bijzonder om op jouw promotie als paranimf naast je te staan. Ik ben blij voor je dat je de voor jou perfecte baan gevonden hebt. **Janet**, ook al leerden we elkaar pas later kennen, we hebben toch al behoorlijk veel momenten gedeeld. Het was erg fijn met je te kunnen sparren over onderzoeksideeën en studieprotocollen. Daarnaast vind ik jouw doorzettingsvermogen bewonderenswaardig. Ik ben er dan ook van overtuigd dat je nog veel mooie dingen gaat bereiken. **Illja**, je bent een hele tijd mijn kamergenootje geweest. Ga ik ooit nog een kamergenootje krijgen die net zo fanatiek als ik sportwedstrijden probeert te volgen met haperende wifi in een oranje versierde kamer? Jij bent inmiddels ook bijna klaar met je proefschrift. Heel veel succes met de laatste loodjes! **Sandra**, dit laatste geldt ook voor jou. Ik heb enorme bewondering voor de manier waarop jij afgelopen jaren je gezin, reisafstand en werk gecombineerd hebt. Hopelijk kan ook jij binnenkort je onderzoeksperiode

met succes afsluiten. **Corrie**, mijn favoriete lunchmaatje. We hebben samen heel wat uurtjes doorgebracht in het restaurant. Ik heb ook ontzettend gelachen om sommige uitspraken van je. 'TM!' ☺. Dank aan de **arts-assistenten** op de afdeling voor de gezellige lunches en uitjes. Gelukkig gaan we elkaar nog regelmatig tegenkomen tijdens het maandelijkse AIOS-onderwijs. Tot slot wil ik de **vele collega's** bedanken die ervoor hebben gezorgd dat ik een ontelbare hoeveelheid geslaagde vrijdagmiddag-borrels heb gehad in de **Aesculaaf**. Deze borrels, met aansluitend vaak een etentje in de stad en soms nog wat meer borrels erna, zijn een belangrijk onderdeel van de fantastische tijd die ik tijdens mijn onderzoek heb gehad.

Nieuwe fase: Maastricht!

Ook al vond ik mijn onderzoekstijd fantastisch, uiteindelijk had ik veel zin om weer aan de slag te gaan in de kliniek. Voor mij blijft het patiëntencontact toch een heel erg mooi onderdeel van mijn vak. Met name richting het einde van mijn onderzoekstijd, tijdens het dagenlange schrijven van artikelen en mijn proefschrift, ging ik dit patiëntencontact erg missen. Gelukkig kon ik aansluitend aan mijn PhD traject aan de slag op de afdeling klinische genetica in Maastricht. Het heeft mij verbaasd hoe snel ik gewend was aan mijn 'nieuwe' leven. Veel dank aan al mijn **Maastrichtse collega's** die ervoor hebben gezorgd dat ik mij zo snel op mijn plek heb gevoeld. Ik heb heel veel zin om aankomende jaren mijn opleiding met jullie op de afdeling te doorlopen. **Yvonne**, dank je wel dat je mij de ruimte gaf om eerst mijn proefschrift af te maken, alvorens te starten aan mijn opleiding. Dankzij jouw vooruitziende blik heb ik in alle rust kunnen starten met een nieuwe fase, waar ik nu erg blij mee ben.

Maaïke, tijdens mijn onderzoekstijd ben ik al weleens bij je langs geweest om samen een patiënt te zien op de poli. Ik ben niet vergeten dat jij me aansluitend vroeg om mee te gaan lunchen met de rest van de Maastrichtse collega's. Hierdoor heb je me een kijkje gegeven in de genetica-afdeling Maastricht, wat er uiteindelijk mede toe heeft geleid dat ik besloten heb in Maastricht aan de slag te willen gaan. **Connie**, ook wij hebben afgelopen jaren regelmatig contact gehad met elkaar. Jouw drive om de nieuwste onderzoeksresultaten te combineren met de beste patiëntenzorg is voor mij inspirerend en motiverend. Ik hoop nog veel van je te kunnen leren aankomende jaren. **Margje**, als ik zag dat een patiënt met een interessante mutatie door jou was ingestuurd voor WES, was ik altijd erg blij. Als er iemand snel was met het aanleveren van complete medische gegevens van patiënten was jij het wel. Als ik nu zie hoeveel werk je verricht op de afdeling kan ik hier alleen nog maar meer respect voor hebben. Mijn kamergenootjes **Vyne en Vivian**: wat hebben we het gezellig met elkaar. We hebben nog heel wat leuke jaren voor de boeg. **Encarna**, nu nog collega-AIOS, maar over een paar weken ben je niet alleen internist, maar ook klinisch geneticus. Ik heb veel respect voor de manier waarop jij keuzes maakt om je uiteindelijke doel te behalen. Tot slot, **Sander**, je was tijdens mijn onderzoeksperiode in Nijmegen altijd al mijn vraagbaak in Maastricht. Dat is nu eigenlijk niet veranderd. Dank je wel voor je enthousiasme en interesse voor mijn proefschrift!

Er is meer dan werk alleen

Vers van de schoolbanken begon ik aan mijn eerste ‘echte’ baan. Al snel werd voor mezelf duidelijk dat ik het belangrijk vond om naast het werk tijd te hebben om hier juist helemaal los van te komen. Terugkijkend denk ik dat het vrij aardig gelukt is, ondanks de hoge werkdruk die het publiceren van artikelen, het bezoeken van congressen en het geven van presentaties zo nu en dan met zich mee bracht. **‘Norbertusmeiden’**, jullie ken ik al zo ontzettend lang. Ondanks dat we over het hele land verspreid wonen, zien we elkaar nog steeds met regelmaat en hebben we inmiddels heel wat bijzondere momenten samen gevierd. **Mirthe**, ik vind het heel bijzonder dat jij straks als paranimf naast me staat. We zien elkaar niet wekelijks, maar het is iedere keer weer alsof het een dag geleden is dat ik je heb gesproken. **Ike**, we hebben de afgelopen jaren heel wat kopjes thee gedronken. Ondanks dat we nu wat minder vaak samen thee kunnen drinken, vind ik het ontzettend fijn voor je dat je, samen met Muhammad, je plek hebt gevonden in Dordrecht. **Tobias en Deborah**, het is altijd weer gezellig om samen met Michiel bij jullie en Abel (en binnenkort zijn broertje!) in Scheveningen langs te komen. **Christian**, officieel ken ik je natuurlijk via het werk, maar inmiddels zie ik je ook regelmatig daarbuiten. We delen een passie voor muziek – dat ontdekten we wel tijdens onze lange gesprekken op vrijdagavond in café Samson. We hebben inmiddels al behoorlijk wat concerten bezocht, ik hoop dat er nog velen gaan volgen. **Jasmijn**, het is zo leuk dat we contact zijn blijven houden na ons avontuur in Indonesië! We hebben elkaar daar heel goed leren kennen en daarom is het iedere keer weer gezellig om samen een van onze hobby’s uit te voeren: lekker eten. Tijdens mijn onderzoekstijd ben ik naast werk veel bezig geweest met mijn grootste passie: viool spelen. **Inge, Laurens en Mirjam**, en later **Desiree, Martijn en Maaïke**: we hebben toch maar een heel mooi orkest opgericht. Het was af en toe puzzelen om alles te organiseren om onze drukke banen heen, maar het resultaat mag er wezen denk ik. **Joris, Clara, Olga, Hinke, Tom, Emiel, Dirk, Maaïke, Franke, Hester, Sophie, Simeon, Desiree, Loes, Florian, Sylvie, Christine, en alle andere NNK-ers**: we hebben nog samen gespeeld in het studentenorkest. Hoe leuk is het om muziek te kunnen blijven maken met elkaar!

Lot, het was voor mij niet moeilijk om jou te kiezen als paranimf. Vanaf het begin af aan hebben wij lief en leed gedeeld. Je was altijd mijn maatje op het werk, maar ook daarnaast hebben we maar al te vaak leuke dingen ondernomen. Jouw tomeloze enthousiasme werkte maar al te vaak aanstekelijk op mij. En hoe leuk is het dat we in 2023 (ja, echt, zo lang nog!) allebei klinisch geneticus én gepromoveerd zijn 😊. **Ot-Blokhokkers**: via Michiel ben ik onderdeel geworden van jullie bijzondere vriendengroep. Ook al heeft inmiddels iedereen zijn eigen leven, het blijft toch leuk om met zijn allen bij elkaar te komen. **Emiel en Paula**, jullie waren erbij toen ik - notabene tijdens een etentje - mijn allereerste artikel ging submitten. Wat een spanning om 11 uur ’s avonds! Onze etentjes, inmiddels met Ionica erbij, blijven wat mij betreft voor herhaling vatbaar.

Lieve **(schoon)familie**, dank jullie wel voor de interesse die jullie getoond hebben de afgelopen jaren in mijn promotietraject. Ik weet dat ik niet het gemakkelijkste onderwerp heb uitgekozen om op te promoveren, maar met dit boekje hoop ik dat jullie toch een beetje een beeld krijgen van waar ik afgelopen jaren mee bezig ben geweest. Dank jullie wel ook voor alle gezellige uitjes, etentjes en lieve kaartjes zo nu en dan! Lieve **opa en oma**, daar is mijn boekje dan! Er is geen moment voorbijgegaan dat jullie niet vol interesse vroegen hoe het ging met 'de studie'. Het is altijd weer leuk om bij jullie thuis een kopje thee te drinken of een kop gewone soep te eten. Ik vind het heel bijzonder dat jullie erbij zullen zijn op deze speciale dag! **Lieve broers, Roliene en Cas**, als ik terugdenk aan de afgelopen jaren komen er veel leuke herinneringen naar boven. Het lange weekend in Rome, leuke uitstapjes, gezellige spelletjesavonden, borrels en lekkere etentjes. En natuurlijk de mooiste herinnering: de geboorte van de kleine Cas. Ik hoop dat we, net als afgelopen jaren, nog veel mooie dingen zullen meemaken. **Lieve papa en mama**, dank jullie wel dat jullie altijd onvoorwaardelijk achter de keuzes hebben gestaan die ik gemaakt heb. Jullie waren er op momenten dat ik twijfelde en het moeilijk vond om beslissingen te nemen. En jullie waren er ook om alle mooie en leuke momenten samen te vieren. Het is fijn om in mijn achterhoofd de wetenschap te hebben dat ik altijd bij jullie terecht kan. Ik heb veel zin in aankomende jaren, waarin wij samen, met het gezin en de familie hopelijk weer veel gezelligheid en mooie momenten samen zullen hebben. **Liefste Michiel**: mijn werkstukje is klaar. Wat een hoop werk was het, ook voor jou. Gedurende bijna vijf jaar heb jij menig suggesties gedaan over wat er allemaal in mijn dankwoord moest komen. Twee kantjes gericht aan jou zou toch het minimum zijn. Dit was natuurlijk als grap bedoeld. Maar die twee kantjes zou ik met gemak vol schrijven. Het was zo ongelofelijk fijn om met jou te sparren over de grote lijnen van mijn proefschrift, na te denken over figuren en uiteindelijk: de lay-out van mijn proefschrift te maken. Jouw frisse kijk op mijn onderzoek werkte op cruciale momenten zo verhelderend. Na jouw academietijd ligt nu ook mijn promotietijd achter ons. Ik heb ontzettend veel zin in de tijd samen die voor ons ligt. Ik ben ervan overtuigd dat we, net als de afgelopen jaren, met veel geluk en plezier samen alle nieuwe uitdagingen aan zullen gaan.



Curriculum Vitae

Margot Reijnders was born on 27th July 1989 in Roosendaal en Nispen (Noord-Brabant, The Netherlands). Currently, she lives in Arnhem with her husband Michiel.



EDUCATION

- (2007) Gymnasium, Norbertuscollege, Roosendaal, The Netherlands.
- (2014) Master of Science Medicine, Radboud University, The Netherlands.
 - * Internship 'Tropical Diseases' in Semarang, Indonesia (2013)



WORK EXPERIENCE

- (2014-2018): Clinical researcher, PhD project entitled 'Neurodevelopmental disorders: a next generation', Radboudumc, Nijmegen, The Netherlands.
- (2018-Current): Clinical geneticist in training, Maastricht UMC+, Maastricht, The Netherlands.



PRESENTATIONS

- (2015) American Society of Human Genetics Annual Meeting 2015, Baltimore, USA. *
- (2015) NVHG Najaarsymposium 2015, Arnhem, The Netherlands. §
- (2016) Joint UK Dutch Genetics Conference, Cardiff, Great-Brittain. *
- (2016) 1st PURA Syndrome Conference 2016, London, Great-Brittain. §
- (2016) The European Human Genetics Conference 2016, Barcelona, Spain. *
- (2016) PURA Syndrome Dutch Parents Gathering, Rosmalen, The Netherlands. §
- (2016) American Society of Human Genetics Annual Meeting 2016, Vancouver, USA. **
- (2016) Mini-symposium Institute of structural Biology, Munich, Germany. §
- (2017) The European Human Genetics Conference 2017, Copenhagen, Denmark. *
- (2017) 2nd PURA Syndrome Conference 2017, Philadelphia, USA. §
- (2017) NVK Jonge onderzoekersdag Erfelijke en Aangeboren aandoeningen, Nijmegen, The Netherlands. §
- (2017) American Society of Human Genetics Annual Meeting 2017, Orlando, USA. *
- (2014) 9th International Meeting on Copy Number Variants and Genes in Intellectual Disability and Autism, Troina, Italy. *
- (2015) The European Human Genetics Conference 2015, Glasgow, Great-Brittain. *

- (2018) 3rd Dutch Neurodevelopmental Disorders meeting, Rotterdam, The Netherlands. §
- (2018) Van Gelderenwerkgroep meeting 2018, Utrecht, The Netherlands. §
- (2018) 3rd PURA Syndrome Conference 2018, Hinxton, Great-Brittain. §
- (2018) 29th European Meeting on Dysmorphology, Bisschofsheim, France.*
- (2015) Manchester Dysmorphology Training course, Manchester, Great-Brittain
- (2016) Popular writing with Kennislink, Radboud University, Nijmegen, The Netherlands
- (2016) 5th Course on Next Generation Sequencing, Bertinoro di Romagna, Italy
- (2017) e-BROK cursus, Nijmegen, The Netherlands

§ Invited presentation

* Oral presentation after abstract selection

** Poster presentation after abstract selection



OTHER



AWARDS

- (2015) ASHG/Charles J. Epstein Trainee Award for Excellence in Human Genetics Research – Semifinalist (£700)
- (2017) ASHG/Charles J. Epstein Trainee Award for Excellence in Human Genetics Research – Semifinalist (£750)
- (2018) Best presentation 29th European Meeting on Dysmorphology 2018 (€800)



COURSES

- (2014) Introduction to Data-analysis, Erasmus Summer Programme, Rotterdam, The Netherlands
- (2014) Advanced Conversation, Radboud University, Nijmegen, The Netherlands
- (2014) Goldrain course in Clinical Cytogenetics, Zuid-Tirol, Italy
- (2014) 27th Course in Medical Genetics, Bertinoro di Romagna, Italy
- (2015) Education in a nutshell, Radboud University, Nijmegen, The Netherlands
- Member of the PURA syndrome Medical Advisory Team (www.purasyndrome.org) (2015 - current)
- Clinical position in the 'Registry Advisory Committee' of the PURA Foundation (2016 – current)
- Chairman Neurodevelopmental Disorders Meeting Radboudumc (2016 – 2017)
- Experience with teaching of medical students: leading of working groups, response colleges, individual and group projects (2014 – 2018)
- Co-founder and chairman of the Nieuw Nijmeegs Kamerorkest (2015 – 2017)
- Coordinator of the foundation 'Uitzicht', a voluntary organization for undocumented refugees (2015 – 2018)

List of publications

Reijnders MRF*, Kousi M*, van Woerden GM*, Klein M, Bralten J, Mancini GMS, van Essen T, Proietti-Onori M, Smeets EEJ, van Gastel M, Stegmann APA, Stevens SJC, Lelieveld SH, Gilissen C, Pfundt R, Tan PL, Kleefstra T, Franke B, Elgersma Y*, Katsanis N*, Brunner HG*. Variation in a range of mTOR-related genes associates with intracranial volume and intellectual disability. [*Nat Commun.* 2017 Okt 20;8(1):1052]

Reijnders MRF*, Zachariadis V*, Latour B*, Jolly L*, Mancini GM, Pfundt R, Wu KM, van Ravenswaaij-Arts CM, Veenstra-Knol HE, Anderlid BM, Wood SA, Cheung SW, Barnicoat A, Probst F, Magoulas P, Brooks AS, Malmgren H, Harila-Saari A, Marcelis CM, Vreeburg M, Hobson E, Sutton VR, Stark Z, Vogt J, Cooper N, Lim JY, Price S, Lai AH, Domingo D, Reversade B; DDD Study, Gecz J, Gilissen C, Brunner HG, Kini U*, Roepman R*, Nordgren A*, Kleefstra T*. De Novo Loss-of-Function Mutations in USP9X Cause a Female-Specific Recognizable Syndrome with Developmental Delay and Congenital Malformations. [*Am J Hum Genet.* 2016 Feb 4;98(2):373-81]

Reijnders MRF*, Ansor NM*, Kousi M*, Yue WW, Tan PL, Clarkson K, Clayton-Smith J, Corning K, Jones JR, Lam WWK, Mancini GMS, Marcelis C, Mohammed S, Pfundt R, Roifman M, Cohn R, Chitayat D, Deciphering Developmental Disorders Study, Millard TH, Katsanis N*, Brunner HG*, Banka S*. RAC1 Missense Mutations in Developmental Disorders with Diverse Phenotypes. [*Am J Hum Genet.* 2017 Sep 7;101(3):466-477]

Reijnders MRF*, Janowski R, Alvi M, Self JE, van Essen TJ, Vreeburg M, Rouhl RPW, Stevens SJC, Stegmann APA, Schieving J, Pfundt R, van Dijk K, Smeets E, Stumpel CTRM, Bok LA, Cobben JM, Engelen M, Mansour S, Whiteford M, Chandler KE, Douzgou S, Cooper NS, Tan EC, Foo R, Lai AHM, Rankin J, Green A, Lönnqvist T, Isohanni P, Williams S, Ruhoy I, Carvalho KS, Dowling JJ, Lev DL, Sterbova K, Lassuthova P, Neupauerová J, Waugh JL, Keros S, Clayton-Smith J, Smithson SF, Brunner HG, van Hoeckel C, Anderson M, Clowes VE, Siu VM, DDD Study T, Selber P, Leventer RJ, Nellaker C, Niessing D, Hunt D*, Baralle D*. PURA syndrome: clinical delineation and genotype-phenotype study in 32 individuals with review of published literature. [*J Med Genet.* 2018 Feb;55(2):104-113]

Reijnders MRF*, Miller KA*, Alvi M, Goos JAC, Lees MM, de Burca A, Henderson A, Kraus A, Mikat B, de Vries BBA, Isidor B, Kerr B, Marcelis C, Schluth-Bolard C, Deshpande C, Ruivenkamp CAL, Wiczorek D; Deciphering Developmental Disorders Study, Baralle D, Blair EM, Engels H, Lüdecke HJ, Eason J, Santen GWE, Clayton-Smith J, Chandler K, Tatton-Brown K, Payne K, Helbig K, Radtke K, Nugent KM, Cremer K, Strom TM, Bird LM, Sinnema M, Bitner-Glindzicz M, van Dooren MF, Alders M, Koopmans M, Brick L, Kozenko M, Harline ML, Klaassens M, Steinraths M, Cooper NS, Edery P, Yap P, Terhal PA, van der Spek PJ, Lakeman P, Taylor RL, Littlejohn RO, Pfundt R, Mercimek-Andrews

S, Stegmann APA, Kant SG, McLean S, Joss S, Swagemakers SMA, Douzgou S, Wall SA, Kürty S, Calpena E, Koelling N, McGowan SJ, Twigg SRF, Mathijssen IMJ, Nellaker C, Brunner HG*, Wilkie AOM*. De Novo and Inherited Loss-of-Function Variants in TLK2: Clinical and Genotype-Phenotype Evaluation of a Distinct Neurodevelopmental Disorder. *[Am J Hum Genet. 2018 Jun 7;102(6):1195-1203]*

Reijnders MRF, Leventer RJ, Lee BH, Baralle D^{4,5}, Selber P⁶, Paciorkowski AR³, Hunt D⁴. (2017) PURA-Related Neurodevelopmental Disorders. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® *[Internet]*. Seattle (WA): University of Washington, Seattle; 1993-2018.

Stefan H. Lelieveld*, **Margot R.F. Reijnders***, Rolph Pfundt, Helger G. Yntema, Erik-Jan Kamsteeg, Petra de Vries, Bert. B.A. de Vries, Marjolein H. Willemsen, Tjitske Kleefstra, Katharina Löhner, Maaïke Vreeburg, Servi Stevens, Ineke van der Burgt, Ernie M.H.F. Bongers, Alexander P.A. Stegmann, Patrick Rump, Tuula Rinne, Marcel R. Nelen, Joris A. Veltman, Lisenka E.L.M. Vissers*, Han G. Brunner*, Christian Gilissen*. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *[Nat Neurosci. 2016 Sep;19(9):1194-6]*

Ideke J.C. Lamers*, **Margot R.F. Reijnders***, Hanka Venselaar, Alison Kraus, DDD Study, Sandra Jansen, Bert B.A. de Vries, Gunnar Houge, Gyri Aasland Gradek, Jieun Seo, Murim Choi, Jong- Hee Chae, Ineke van der Burgt, Rolph Pfundt, Stef J.F. Letteboer, Sylvia E.C. van Beersum, Simone Dusseljee, Han G. Brunner, Dan Doherty, Tjitske Kleefstra*, Ronald Roepman*. Recurrent de novo heterozygous mutations disturbing the GTP/GDP binding pocket of RAB11B cause intellectual disability and a distinctive brain phenotype. *[Am J Hum Genet. 2017 Nov 2;101(5):824-832]*

D. Lugtenberg*, **M.R.F. Reijnders***, M. Fenckova*, E.K. Bijlsma, R. Bernier, B. W.M. van Bon, E. Smeets, A.T. Vulto-van Silfhout, D. Bosch, E.E. Eichler, H.C. Mefford, G.L. Carvill, E.M.H.F. Bongers, J.H.M. Schuurs-Hoeijmakers, C.A. Ruivenkamp, G.W.E. Santen, A.M.J.M. van den Maagdenberg, C.M.P.C.D. Peeters-Scholte, S. Kuenen, P. Verstreken, R. Pfundt, H.G. Yntema, P.F. de Vries, J.A. Veltman, A. Hoischen, C. Gilissen, B.B.A. de Vries, A. Schenck*, T. Kleefstra*, L.E.L.M. Vissers*. De novo loss-of function mutations in WAC cause a recognizable intellectual disability syndrome and are associated with learning deficits in Drosophila. *[Eur J Hum Genet. 2016 Aug;24(8):1145-53]*

Jung-Hyun Kim*, Deepali N. Shinde*, **Margot R.F. Reijnders***, Natalie S Hauser, Rebecca L Belmonte, Gregory R Wilson, Daniëlle G. Bosch, Paula A. Bubulya, Vandana Shashi, Slavé Petrovski, Joshua K. Stone, Eun Young Park, Joris A. Veltman, Margje Sinnema, Connie T.R.M. Stumpel, Jos M. Draaisma, Joost Nicolai, University of Washington Center for Mendelian Genomics, Helger G. Yntema, Kristin Lindstrom, Bert BA de Vries, Tamison Jewett, Stephanie L. Santoro, Julie Vogt, The Deciphering Developmental Disorders Study, Kristine K. Bachman, Andrea H. Seeley, Alyson Krokosky, Clessen Turner, Luis Rohena, Maja Hempel, Fanny Kortüm, Davor Lessel, Axel Neu, Tim M. Strom, Dagmar Wieczorek, Nuria Bramswig, Franco A. Laccone, Jana Behunova, Helga Rehder, Christopher T. Gordon, Marlène Rio, Serge Romana, Sha Tang, Dima El-Khechen, Megan T. Cho, Kirsty McWalter, Ganka

Douglas, Berivan Baskin, Amber Begtrup, Tara Funari, Kelly Schoch, Alexander P.A. Stegmann, Servi J.C. Stevens, Dong-Er Zhang, David Traver, Xu Yao, Daniel G. MacArthur, Han G. Brunner, Grazia M. Mancini, Richard M. Myers, Laurie B. Owen, Ssang-Taek Lim, David L. Stachura, Lisenka E.L.M. Vissers*, Eun-Young Erin Ahn*. De novo mutations in SON disrupt RNA splicing of genes essential for brain development and metabolism, causing an intellectual-disability syndrome. *[Am J Hum Genet. 2016 Sep 1;99(3):711-719]*

Ba W*, Yan Y*, **Reijnders MRF***, Schuurs-Hoeijmakers JH, Feenstra I, Bongers EM, Bosch DG, De Leeuw N, Pfundt R, Gilissen C, De Vries PF, Veltman JA, Hoischen A, Mefford HC, Eichler EE, Vissers LE, Nadif Kasri N*, De Vries BB*. TRIO loss of function is associated with mild intellectual disability and affects dendritic branching and synapse function. *[Hum Mol Genet. 2016 Mar 1;25(5):892-902]*

Lee BH, **Reijnders MRF**, Abubakare O, Tuttle E, Lape B, Minks KQ, Stodgell C, Bennetto L, Kwon J, Fong CT, Gripp KW, Marsh ED, Smith WE, Huq AM, Coury SA, Tan WH, Solis O, Mehta RI, Leventer RJ, Baralle D, Hunt D, Paciorkowski AR. Expanding the neurodevelopmental phenotype of PURA syndrome. *[Clin Genet. 2018 Apr;93(4):880-890]*

Chiu ATG, Pei SLC, Mak CCY, Leung GKC, Yu MHC, Lee SL, Vreeburg M, Pfundt R, van der Burgt I, Kleefstra T, Frederic TM, Nambot S, Faivre L, Bruel AL, Rossi M, Isidor B, Küry S, Cogne B, Besnard T, Willems M, **Reijnders MRF**, Chung BHY. Okur-Chung neurodevelopmental syndrome: Eight additional cases with implications on phenotype and genotype expansion. *[Clin Genet. 2018 Apr;93(4):880-890]*

Snijders Blok L*, Madsen E*, Juusola J*, Gilissen C, Baralle D, **Reijnders MRF**, Venselaar H, Helsmoortel C, Cho MT, Hoischen A, Vissers LE, Koemans TS, Wissink-Lindhout W, Eichler EE, Romano C, Van Esch H, Stumpel C, Vreeburg M, Smeets E, Oberndorff K, van Bon BW, Shaw M, Gecz J, Haan E, Bienek M, Jensen C, Loeys BL, Van Dijk A, Innes AM, Racher H, Vermeer S, Di Donato N, Rump A, Tatton-Brown K, Parker MJ, Henderson A, Lynch SA, Fryer A, Ross A, Vasudevan P, Kini U, Newbury-Ecob R, Chandler K, Male A; DDD Study, Dijkstra S, Schieving J, Giltay J, van Gassen KL, Schuurs-Hoeijmakers J, Tan PL, Pediaditakis I, Haas SA, Retterer K, Reed P, Monaghan KG, Haverfield E, Natowicz M, Myers A, Kruer MC, Stein Q, Strauss KA, Brigatti KW, Keating K, Burton BK, Kim KH, Charrow J, Norman J, Foster-Barber A, Kline AD, Kimball A, Zackai E, Harr M, Fox J, McLaughlin J, Lindstrom K, Haude KM, van Roozendaal K, Brunner H, Chung WK, Kooy RF, Pfundt R, Kalscheuer V, Mehta SG*, Katsanis N*, Kleefstra T*. Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *[Am J Hum Genet. 2015 Aug 6;97(2):343-52]*

Koemans TS, Kleefstra T, Chubak MC, Stone MH, **Reijnders MRF**, de Munnik S, Willemsen MH, Fenckova M, Stumpel CTRM, Bok LA, Sifuentes Saenz M, Byerly KA, Baughn LB, Stegmann APA, Pfundt R, Zhou H, van Bokhoven H, Schenck A, Kramer JM. Functional convergence of histone methyltransferases EHMT1 and KMT2C involved in intellectual disability and autism spectrum disorder. *[PLoS Genet. 2017 Oct 25;13(10):e1006864]*

Olson HE, Jean-Marçais N, Yang E, Heron D, Tatton-Brown K, van der Zwaag PA, Bijlsma EK, Krock BL, Backer E, Kamsteeg EJ, Sinnema M, **Reijnders MRF**, Bearden D, Begtrup A, Telegrafi A, Lunsing RJ, Burglen L, Lesca G, Cho MT, Smith LA, Sheidley BR, El Achkar CM, Pearl PL, Poduri A, Skraban CM, Tarpinian J, Nesbitt AI, Fransen van de Putte DE, Ruivenkamp CAL, Rump P, Chatron N, Sabatier I, De Bellescize J, Guibaud L, Sweetser DA, Waxler JL, Wierenga KJ; DDD Study, Donadieu J, Narayanan V, Ramsey KM; C4RCD Research Group, Nava C, Rivière JB, Vitobello A, Mau-Them FT, Philippe C, Bruel AL, Duffourd Y, Thomas L, Lelieveld SH, Schuurs-Hoeijmakers J, Brunner HG, Keren B, Thevenon J, Faivre L, Thomas G, Thauvin-Robinet C. A Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis. [*Am J Hum Genet.* 2018 Oct 4;103(4):631]

Lessel D, Schob C, Küry S, **Reijnders MRF**, Harel T, Eldomery MK, Coban-Akdemir Z, Denecke J, Edvardson S, Colin E, Stegmann APA, Gerkes EH, Tessarech M, Bonneau D, Barth M, Besnard T, Cogné B, Revah-Politi A, Strom TM, Rosenfeld JA, Yang Y, Posey JE, Immken L, Oundjian N, Helbig KL, Meeks N, Zegar K, Morton J, The Ddd Study, Schieving JH, Claasen A, Huentelman M, Narayanan V, Ramsey K; C4RCD Research Group, Brunner HG, Elpeleg O, Mercier S, Béziau S, Kubisch C, Kleefstra T, Kindler S, Lupski JR, Kreienkamp HJ. De Novo Missense Mutations in DHX30 Impair Global Translation and Cause a Neurodevelopmental Disorder. [*Am J Hum Genet.* 2018 Jan 4;102(1):196]

Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, Friend KL, Baker C, Buono S, Vissers LE, Schuurs-Hoeijmakers JH, Hoischen A, Pfundt R, Krumm N, Carvill GL, Li D, Amaral D, Brown N, Lockhart PJ, Scheffer IE, Alberti A, Shaw M, Pettinato R, Tervo R, de Leeuw N, **Reijnders MR**, Torchia BS, Peeters H, O'Roak BJ, Fichera M, Hehir-Kwa JY, Shendure J, Mefford HC, Haan E, Géczy J, de Vries BB, Romano C, Eichler EE. Refining analyses of copy number variation identifies specific genes associated with developmental delay. [*Nat Genet.* 2014 Oct;46(10):1063-71]

Bosch DG, Boonstra FN, **Reijnders MR**, Pfundt R, Cremers FP, de Vries BB. Chromosomal aberrations in cerebral visual impairment. [*Eur J Paediatr Neurol.* 2014 Nov;18(6):677-84]

Houge G, Haesen D, Vissers LE, Mehta S, Parker MJ, Wright M, Vogt J, McKee S, Tolmie JL, Cordeiro N, Kleefstra T, Willemsen MH, **Reijnders MR**, Berland S, Hayman E, Lahat E, Brilstra EH, van Gassen KL, Zonneveld-Huijssoon E, de Bie CL, Hoischen A, Eichler EE, Holdhus R, Steen VM, Døskeland SO, Hurles ME, FitzPatrick DR, Janssens V. B56δ-related protein phosphatase 2A dysfunction identified in patients with intellectual disability. [*J Clin Invest.* 2015 Aug 3;125(8):3051-62]

* Shared authorship

