UNIVERSITA' DEGLI STUDI DI TORINO

PhD school in Biomedical Sciences and Oncology

Human Genetics Curriculum

XXXIV study cycle

UNIVERSITÀ DI TORINO

NEUROWES PROJECT: FROM WHOLE EXOME SEQUENCING DATA ANALYSIS TO FUNCTIONAL CHARACTERIZATION OF NOVEL NDD-ASSOCIATED GENES

PhD thesis

Director
Professor Emilio Hirsch

Supervisor
Professor Alfredo Brusco

Candidate

Lisa PAVINATO

Academic year 2021/2022

Summary

Introduction

Neurodevelopmental disorders (NDDs) affect more than 3% of population worldwide[1], and include autism spectrum disorder (ASD), intellectual disability (ID), attention deficit hyperactivity disorder (ADHD), and epilepsy[2]. They are distinguished by high clinical heterogeneity, with impaired cognition, communication, adaptive behaviour, and psychomotor delay.

For a long time, first-tier tests for patients with NDDs have been G band karyotype, *FMR1* trinucleotide repeats analysis and chromosomal microarray, with a diagnostic rate of about 15%[3,4].

The advent of next generation sequencing techniques (NGS) led to a significant increase in the understanding of the genetic causes of NDDs, identifying more than 900 genes as involved in pathogenic processes and associating more than one hundred of them to a monogenic form of NDD (https://gene.sfari.org/)[5,6]. To date, the most used technique is exome sequencing (ES), which can lead to a diagnostic yield up to 40% when the proband and both the biological parents are considered[2].

The identification of the genetic aetiology of the disease has a fundamental importance for the genetic counselling and patient management, as it can determine a better assessment of the reproductive risk and gives the possibility to establish the best clinical care and medication for the patient. Despite this and the enormous advances done in the field of the genetics of NDDs, many individuals do not receive a molecular diagnosis.

Several reasons can be identified for these missing diagnoses. First, pathogenesis of NDDs is highly heterogeneous, with elevate level of complexity and the contribution of both genetics and non-genetics factors. Several studies are highlighting two essential principles that can clarify the genetics of NDDs: the gene vulnerability and the mutational load[2,7].

Gene vulnerability can be defined as the capability of a given gene to tolerate disruptive variants. To date, several genes have been reported as highly intolerant to variation[8] and more than 4,500 genes have been shown to be highly sensitive to dosage[9]. Variants in those genes are more likely to be subject of a strong negative selective pressure, and result in rare monogenic forms of NDDs, in which a major contributor to risk is a single gene mutation or copy number variation (CNV)[10,11].

The other end of the susceptibility spectrum includes genes with low sensitivity to disruptive mutations, in which variants are more likely to be transmitted among generation, due to the lack of negative selective pressure[10]. Those variants are generally common in the population, and their pathological effect is shown only when the sum of thousands of common risk alleles

with small effects pass a certain level[10,12–14]. The contribution of all common variants to a trait can be summarized as a polygenic risk score (PRS)[15].

Along this spectrum -from monogenic to polygenic forms of NDDs-, there are other several possible mechanisms, including dual molecular diagnosis and intra- and extragenic two-hit models[2]. Moreover, also other risk factors should be considered: advanced parental age, birth trauma and valproate exposition during pregnancy has been shown to contribute to NDDs onset[16], as well as other complex genetic and environmental regulation, including epigenetic mechanisms[17,18].

In addition to the causal complexity, NDDs are represented also by a high phenotypic heterogeneity. This can arise from multiple sources, including differences in core diagnostic features, sex, age, and co-occurring conditions. Furthermore, the lack of standardize clinical reports worldwide, further increases the level of intricacy[19]. In many cases, even variants in the same gene can be associated to a broad spectrum of phenotypes: an example is the gene *SCN2A*, for which several distinct phenotypes have been reported, ranging from early infantile epilepsies to later-onset seizures and ASD/ID without epilepsy[20].

Our effort in understanding the genetics of NDDs started with the NeuroWES project, an international collaborative work began in 2015. Our aim was to collect and NDD families and use trio exome sequencing (trio-ES) as diagnostic/research tool. A brief dissertation of the project will be proposed in Chapter 1.

The analysis of trio-ES data has been rarely a straightforward genotype-phenotype correlation, leading to raising of several projects from the NeuroWES, discussed in chapters 2-4.

Chapter 2 will focus on two novel genes we associated with NDDs, namely *CAPRIN1* and *RPH3A*. *CAPRIN1* was found mutated in more than ten patients with a clinical phenotype represented characterized by language impairment/speech delay, ID, ADHD, and ASD. To demonstrate that loss-of-function of one *CAPRIN1* allele is pathogenic, we performed several experiments on cortical neurons derived from human induced pluripotent stem cells (hiPSCs), engineered with CRISPR/Cas9[21].

RPH3A was found mutated in five patients characterized by ID and ASD or ID and seizures. To understand the pathological mechanisms, we performed functional experiments on rat hippocampal neurons, providing evidence of a likely gain-of-function mechanism for the identified variants.

The third chapter will focus on the study of variants of *TLK2*, a gene associated with a NDD in 2018[22]. Despite more than 40 patients described, evidence about the pathological mechanism leading to the disorder were missing. We therefore performed experiments both on patient-

derived cells and using overexpressing models, showing that the analysed variants caused a defect in the kinase activity and in chromatin remodelling[23]. Following this, we also provided preliminary data about a private methylation profile identifiable in *TLK2*-patients.

Chapter 4 focuses on the identification of novel phenotypes associated to known genes. In our cohort, we frequently found variants in genes already associated with a phenotype that however could not explain partially or totally the phenotype observed in our patient. Among the examples, we show that *SUPT5H,* currently associated with a beta-thalassemia trait, can also be responsible for NDD.

Supplementary data are reported in *Appendix: supplementary data* and references are gathered at end of the thesis.

Chapter 1: The NeuroWES project

Introduction and general overview

NeuroWES project started in 2015, in collaboration with the Autism Sequencing Consortium (ASC), an international group of scientists working on understanding the genetic causes of NDDs[24].

Our group collaborated to ASC studies providing a cohort of cases recruited in different national hospitals (*see below*). Currently, we recruited 509 families: 4 singleton, 17 duo, 429 trios, 50 quad, 8 penta, and one esa. All families had at least one affected proband, and in some cases (quad, penta and esa families) further affected or healthy siblings were available.

The workflow of processing and analysis involved several working groups:

- Recruitment of cases (main Centers): AOU Città della Salute e della Scienza, Turin, Italy; A.O. Santa Croce e Carle, Cuneo, Italy; Policlinico P. Giaccone, Palermo, Italy.
- Sample preparation and shipment: Medical Genetics and Rare Disease laboratory, University of Turin, Turin, Italy.
- Exome sequencing: the Autism Sequencing Consortium (ASC) working group, Broad Institute, New York, USA.
- Variant prioritization: Ospedale Pediatrico Bambino Gesù, Rome; Ospedale Universitario Sant'Orsola-Malpighi, Bologna.
- Data analysis, variant validation, and functional characterization: Medical Genetics and Rare Disease laboratory, University of Turin, Turin, Italy.

For each family, we collected peripheral blood DNA and clinical information as HPO terms[25]. Sequencing of all families' members was performed in collaboration with Professor J. D. Buxbaum and Dr S. De Rubeis at the Broad Institute (Cambridge, MA) on the Illumina HiSeq sequencers as previously described[6,26]. Raw data (.bam or .cram files) were processed by Dr T. Pippucci and A. Bruselles at Policlinico Sant'Orsola-Malpighi (Bologna, Italy) and Bambino Gesù Hospital (Roma, Italy), using in-house implemented pipelines of analysis[27,28] based on the GATK Best Practices[29]. The UCSC GRCh37/hg19 version of genome assembly was used as a reference for reads alignment by means of BWA-MEM[30] tool and the subsequent variant calling with HaplotypeCaller (GATK v3.7). SnpEff v.4.3[31] and dbNSFP v.3.5[32] tools were used for variants functional annotation, including Combined Annotation Dependent Depletion (CADD) v.1.3[33], Mendelian Clinically Applicable Pathogenicity (M-CAP) v.1.0 and Intervar v.0.1.6 for functional impact prediction[34]. Thereby, the analysis was narrowed to variants which affect coding sequences or splice site regions. Moreover, high-quality variants were filtered against public databases (dbSNP150 and gnomAD ver.2.0.1) so that only variants with unknown frequency or

having MAF <0.1%, as well as variants occurring with frequency <1% in internal population-matched database (~2,000 exomes) were considered. To check the accuracy of the reported pedigree information, relatedness was calculated between each pair of samples using HapMap method. Combined with the imputed sex, these inferred pedigrees were compared to reported pedigrees and verified for discrepancies. A further variant stratification in conformity with the American College of Medical Genetics and Genomics (ACMG) guidelines[35], also considering mode of inheritance and functional *in silico* prediction of impact, allowed us considering the final set of variants for possible associations with the phenotype.

In Torino, we analysed prioritized variants within each family, interpreted them, and performed phenotype-genotype correlations. Among the 509 families sent for the analysis, 36 are still to be sequenced, 54 are sequenced, but the analysis is pending, 52 are under clinical evaluation and 13 failed due to technical issues of sequencing. I personally analysed 219 of the 354 families, whose data were available. Overall results showed:

- 120 solved families (34%), with variants in genes associated with NDDs reported in OMIM and classified as likely pathogenic or pathogenic[34];
- 22 families (6%) with variants of uncertain significance (VUS) in genes associated with NDDs in OMIM;
- 173 unsolved families (49%), for which we were not able to identify any causative or likely causative variant;
- 39 families (11%) for which we identified VUS, likely pathogenic or pathogenic variants in genes not reported in OMIM but with literature data supporting their involvement in NDDs.

After the identification of likely pathogenic or pathogenic variants, patients follow a diagnostic workflow, which requires the validation of the variant by Sanger sequencing on a second DNA extraction and a clinical review. Patients with a VUS follow a similar procedure; the uncertain nature of the variant is always clearly stated in the referral, and its status can be reviewed in the following years, depending on eventual re-classification of the variant.

Among the 40% of solved patients (VUS, likely pathogenic and pathogenic variants), we identified

- 50 cases with an X-linked disorder; 6 of them were females;
- 41 cases with an autosomal dominant disorder: 39 of them with a *de novo* variant, one maternal and one of unknown origin;
- 50 cases with an autosomal recessive disorder: 2 with homozygous variants and the remaining with compound heterozygous inheritance;

- 1 case with an imprinting disorder.

In the following paragraphs, I will discuss about peculiar circumstances that occurred in my experience with the analysis of the NeuroWES cohort of cases, from atypical clinical presentations to methods to clarify unsolved cases.

Lessons from the NeuroWES experience: when available literature cannot explain the observed clinical phenotype.

Each patient analysed by trio-ES has its own peculiarities, and sometimes minor clinical details can be lost in the report of HPO terms. Despite the need to be concise and to use well defined clinical terms (as HPO), our experience showed the importance of having very precise clinical description, possibly completed by neuroradiological findings, facial pictures, and an in-depth familial history. Indeed, it happens quite often that the HPO terms reported in our registries cannot explain the whole complexity of the index case and its family, making complex the evaluation of the molecular findings. Thanks to the collaboration with clinicians able to access detailed clinical information, we could often re-evaluate the genetic data of patients in the light of a better description of their familial history and phenotype. The re-evaluation usually has given us two results: i) the identification of a second gene (dual diagnoses) that could explain the additional phenotypic feature and ii) the expansion of the clinical phenotype reported in the literature.

The identification of dual diagnoses in 26 solved cases (18%), highlighted the importance of considering multiple molecular diagnoses even when the first identified genetic hit can explain most of the phenotype. Nevertheless, sometimes clinical features associated with different genes overlap, making problematic to realize if we are observing a combined effect.

The expansion of a clinical phenotype is quite common, especially for genetic disorders reported in a limited number of patients. To this regard, the collection of in-depth clinical information allowed us to add further clinical features to those already known, such as in the case of the Skraban-Deardorff Syndrome, associated with mutations in *WDR26*[36]. We also described novel phenotypes associated with known genes, such as the *PI4KA* gene[37]. This topic will be better discussed in Chapter 4.

Approaches to clarify the effect of variant of uncertain significance.

The impact of variant of uncertain significance is always difficult to understand, as they are generally not “strong” enough to be considered likely pathogenic and not “weak” enough to be

discarded. The use of *in silico* tools can help us to predict the effect uncertain variants, but further data can re-classify of the variant versus the likely benign or likely pathogenic state[35].

To comprehend the effect of a VUS, the optimal approach is performing a functional test, which can test the effect of the variant directly on the protein. In some cases, we took advice from experts on specific genes. As an example, the variant c.354T>A; p.(Asn118Lys) in *SLC9A6* found in a male case with NDD. This gene is generally associated with an X-linked form of syndromic intellectual disability and epilepsy (OMIM #300243). The variant was considered a VUS by *in silico* tools; patient's clinical features seemed different and less severe than those reported in literature, but we considered a possible variable expressivity. We contacted Professor Eric M. Morrow (Brown University, Providence, RI), who could perform a functional test for SLC9A6, introducing our variant into the protein. He verified that the variant did not change this channel function (data not shown), allowing us to re-classify it as likely benign.

To clarify the impact of variants -including VUS- we also used methylation profiling, also called episignature, in collaboration with Dr B. Sadikovic (London Health Science Center, Canada). Briefly, several neurodevelopmental disorders have been associated with a recognizable methylation profile, which can be identified directly on patient's DNA derived from peripheral blood[38,39]. Thanks to this collaboration, we are analysing every patient with variants in gene associated with a recognizable episignature, to confirm or eventually re-classify the functional effect of specific variants.

Currently, we sent 85 DNA samples for the analysis, with syndromes and/or at least one variant in genes associated with an episignature[39]. Of the 27 already analysed samples, we confirmed the molecular diagnosis for 21 (Table 1, green), while for six of them the diagnosis was negative or different from the previous one (Table 1, light orange). In details, the episignature was negative for cases 140556 and 140558 -siblings previously diagnosed with Nicolaides-Baraitser syndrome-, case 190941 -previously classified as a case with Coffin-Siris syndrome 4- and patient FS0208013_65 – previously reported as a case affected by CHARGE Syndrome. Further investigations are ongoing to understand these discrepancies.

Case 150163 was previously classified as a case of Coffin-Syris Syndrome 1 (CSS1), associated with variants in *ARID1B*: however, the clinics was not completely overlapping with the one described in literature and the episignature analysis showed that the methylation profile of the patient was not associated with CSS1 but with Cornelia de Lange Syndrome (CdLS). To date, we still have not found a molecular diagnosis for the CdLS, and we are therefore evaluating whole genome sequencing (WGS) and/or transcriptome analysis for this patient (*see paragraph "Solving unsolved cases: what's next?"*).

Case GM190084 was used to generate a novel episignature associated with *TLK2* gene, but did not cluster with other samples, probably due to the co-existence of a pathogenic variant in *DYNC1H1* (*see Chapter 3* for further details).

For cases 130091, 131749 and 141444, we found two molecular causes of disease (Table 1, white), with one of the two associated genes (*DYNC1I2*, *NBAS* and *RS1* respectively) still not associated with an episignature. For those cases only the genetic cause already associated with a detectable methylation profile has been confirmed and no peculiar profiles were observed, suggesting that i) the second gene does not cause changes in methylation or that ii) the methylation profile associated with the first gene is still solid even if a second one is present.

In total:

- Six out of 27 cases had variants of uncertain significance, of which 3 out 6 (cases 160708, 121116 and 121886) could be re-classified as likely pathogenic variants, while one (case 150163) should be down-classified as a likely benign variant and reviewed;
- Ten out of 27 cases had likely pathogenic variants, of which 5 out of 10 (cases GM110562, 121061, 130091, 141444 and GM182051) can be re-classified as pathogenic, while 4 (cases 140556, 140558, GM190084 and GM190941) should be down-classified as VUS and reviewed;
- Fourteen out of 27 had pathogenic variants, of which 13 out of 14 (151102, 151105, 110212, 131361,131749, GM181933, 121623, 111092, BA2012002_07, 160759, 122063, GM184039 and FS0208013_65) have been confirmed, while one (case GM191199) should be re-evaluated.

Table 1. Preliminary results of episignature analysis.

			trio-ES data							aCGH data		
DNA code	Sex	ACMG	Disease	OMIM disease	OMIM gene	Gene	Inheritance	AD/AR/XLD/XLR	Variant	Inheritance	CNV	EpiSign result
GM110562	M	4	AUTS18	# 615032	* 610528	*CHD8*	De novo	AD	c.2025-1G>C; p.(?)	maternal	dup19p13.2	AUTS18
121061	M	4	MRD57	# 618050	* 608439	*TLK2*	De novo	AD	c.1586A>G; p.(Asp529Gly)			MRD57
151102	F	5	MRD57	# 618050	* 608439	*TLK2*	maternal	AD	c.1357G>T; p.(Glu453Ter)			MRD57
151105	M	5	MRD57	# 618050	* 608439	*TLK2*	maternal	AD	c.1357G>T; p.(Glu453Ter)	de novo maternal	12q24.21 20q11.22	MRD57
160708	M	3	CSS1	# 135900	* 614556	*ARID1B*	De novo	AD	c.2480C>T; p.(Ala827Val)			BAFopathy
110212	M	5	RSTS1	# 180849	* 600140	*CREBBP*	De novo	AD	c.3779+1G>A ; p.(?)			RSTS
130091	M	4 3 3	CSS9 NEDMIBA	# 615866 # 618492	* 600898 * 603331	*SOX11* *DYNC1I2*	De novo Comp. Het.	AD AR	c.159G>T; p.(Met53Ile) c.398G>A; p.(Arg133Gln) paternal c.393G>T; p.(Leu131Phe) maternal	maternal	dup7p21.2	CSS9
131361	M	5	KLEFS1	# 610253	* 607001	*EHMT1*	De novo	AD	c.3000del; p.(Asp1001ThrfsTer9)			KLEFS
150163	M	3	CSS1	# 135900	* 614556	*ARID1B*	De novo	AD	c.3220G>A; p.(Asp1074Asn)	maternal	del 1p32.2	CdLS
131749	M	5 3	FLHS SOPH	# 136140 # 614800	* 611421 * 608025	*SRCAP* *NBAS*	De novo Comp. Het.	AD AR	c.7937_7938delTG ; (p.Val2646GlufsTer2) c.1702G>A; p.(Val568Ile) paternal c.3826C>A; (Pro1276Thr) maternal	maternal	del2p16.3	FLHS
141444	M	4 4	KLEFS1 RS1	# 610253 # 312700	* 607001 * 300839	*EHMT1* *RS1*	De novo maternal	AD XLR	c.3332G>C; (p.Cys1111Ser) c.629T>C; p.(Ile210Thr)			KLEFS
GM181933	M	5	KLEFS1	# 610253	* 607001	*EHMT1*	De novo	AD	c.508delC ; p.(Gln170ArgfsTer112)			KLEFS
140556	M	4	NCBRS	# 601358	*600014	*SMARCA2*	Unknown	AD	c.2296C>G; p.(Leu766Val)	maternal	del6p25.1	Negative
140558	M	4	NCBRS	# 601358	*600014	*SMARCA2*	Unknown	AD	c.2296C>G; p.(Leu766Val)			Negative
121623	M	5	KBGS	# 148050	* 611192	*ANKRD11*	De novo	AD	c.439C>T; p.(Gln147Ter)			KBGS
111092	M	5	MRXHF1	# 309580	* 300032	*ATRX*	maternal	XLR	exon 3-4 deletion			ATRX
121116	M	3	MRXSCJ	# 300534	* 314690	*KDM5C*	maternal	XL	c.1204G>A; p.(Asp335Asn)			MRXSCJ
121886	F	3	MRXSCJ	# 300534	* 314690	*KDM5C*	maternal	XL	c.1204G>A; p.(Asp335Asn)			MRXSCJ
BA2012002_07	F	5	KBGS	# 148050	* 611192	*ANKRD11*	De novo	AD	c.211_226+1del; p.(?)			KBGS
160759	F	5	CSS1	# 135900	* 614556	*ARID1B*	De novo	AD	c.3984G>A; p.(Trp1328Ter)			BAFopathy
122063	M	5	NEDAMSS	# 618088	* 611720	*IRF2BPL*	De novo	AD	c.2160del; p.(Glu720Ter)	paternal	dup8q22.3	NEDAMSS
GM182051	M	4	RENS1	# 309500	* 300463	*PQBP1*	De novo	AD	c.233C>A; p.(Pro78Gln)			RENS1
GM184039	F	5	RTS2	#613684	* 602700	*EP300*	De novo	AD	c.3671+5G>C; r.3590_3671del	maternal	dup7p22.3	RSTS
GM190084	F	4 5	MRD57 MRD13	# 618050 # 614563	* 608439 * 600112	*TLK2* *DYNC1H1*	De novo De novo	AD AD	c.1480C>T ; p.(Arg494Trp) c.11182C>T; p.(Arg3728Ter)	paternal	dupXp22.12	Negative
GM190941	M	4	CSS4	#614609	* 603254	*SMARCA4*	De novo	AD	c.3068A>G ; p.Glu1023Gly	maternal	dup3q23	Negative
GM191199	M	5	MALNS	#614753	* 164005	*NFIX*	De novo	AD	c.95delA; p.(Asn32Ter)			Negative
FS0208013_65	M	5	CHARGE	# 214800	* 608892	*CHD7*	De novo	AD	c.6194G>A; p.(Arg2065His)			CHARGE

confirmed
different/negative diagnosis
Episignature not available

Notes: Only patients for which results are available have been reported. Confirmed diagnosis are highlighted in green; different or negative diagnosis are indicated in light orange. All the variants are reported in hg19. ACMG criteria are obtained using Franklin by Genoox browser (https://franklin.genoox.com/) and are updated on 28th November 2022.

AD=Autosomal Dominant; AR= Autosomal Recessive; XLD=X-Linked Dominant; XLR:X-Linked Recessive; Comp. Het.= Compound heterozygous. Disease's acronyms are according to omim.org.

Novel genes involved in NDDs: from identification to functional characterization.

The identification of novel candidate genes for neurodevelopmental disorders represents a relevant purpose of our studies, and an important value for the NeuroWES project. Our workflow includes the search for potential novel genes in unsolved cases or in solved cases for which the genotype-phenotype correlation does not match consistently with the previous reported literature.

We applied a series of rules to select candidate genes. High impact variants based on ACMG criteria were selected; *de novo* variants were prioritized as the most important, especially if nonsense, such as those biallelic for possible recessive genes and those in X-linked genes for males. We also considered GnomAD constraint scores to evaluate if the variant type was consistent with a possible haploinsufficiency or gain of function/dominant negative mechanism.

A critical evaluation of the literature and databases (e.g., OMIM, MGI) to identify the contribution in pathways commonly involved in NDDs onset[2]. An important database, we usually queried was SFARI Gene (www.gene.sfari.org), which is an updated list of genes involved in autism spectrum disorders and NDDs susceptibility. The database provides a score for each gene which describes its potential association with NDD.

Next, we tried to collect other cases with the same gene involved, to compare the phenotypes and highlight possible similarities. To this purpose, we queried several registries, including GeneMatcher[40], MSSNG (www.research.mss.ng), and Genomics England[41] databases. In many cases, we found other patients with similar clinical features, and could include our cases in ongoing studies for the clinical and molecular characterization of novel NDD genes. This was the case for *GLRA2*[42], *RICTOR*, *RNF40*, *GPRASP1*, *SF3B3* and *ALG5* (ongoing studies). In other cases, as the ones reported in Chapter 2, we performed functional analyses, leading the projects.

The functional characterization of novel NDDs genes is a challenge for scientists, as it requires to face with the complexity of human brain development and function, the heterogeneity behind NDDs and the limitations due to the impossibility of using human brain samples in most of cases[43]. Simple cellular models, as immortalized cells and lymphoblastoid cells may be used, however it should be considered that they do not represent the main affected tissue -the brain- and they could not be used for those genes exclusively expressed in the central nervous system. Until recently, animal models played a central role in the study of NDDs, however, they cannot fully recapitulate the human clinical and/or molecular and cellular phenotypes[44]. In recent years, great efforts have been spent to develop also *in vitro* models of NDDs, particularly for the study of the cerebral cortex, which is responsible for many high-level cognitive functions frequently disrupted in NDDs[43].

A major innovation in the field has been the introduction of human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs), which can be differentiated in several cell lineages, including specific or mixed neuronal populations[43]. With the use of different culture media and specific factors, hESCs and hiPSCs can be differentiated in a plethora of two-dimensional (2D) neuronal cells, including dopaminergic, GABAergic, and glutamatergic neurons; furthermore, models of astrocytes, microglia and oligodendrocytes have been developed[43] . These models constitute a very scalable way to study NDDs, allowing to identify functional and morphological defects. Furthermore, they can be used for *in vitro* drug screenings and therapeutic purposes[45].

Limitations of 2D models are mainly represented by the lack of cellular diversity, microenvironment, and interconnection between different cellular types and/or brain regions.

These limitations can partially be overcome using brain organoids, three-dimensional (3D) cell cultures represented by self-organized stem cells in complex structures[43]. Brain organoids can represent single or multiple brain regions[43], and studies showed they accurately resemble the first and second trimesters of the human foetal brain[46,47].

Choosing the right cellular model to functional characterize a novel gene involved in NDDs is not a black or white decision, and different models may be used to pursue the aims of different studies.

Solving unsolved cases: what's next?

NeuroWES project laid the foundations for many future projects in the next future; one important challenge will be solving the "unsolved cases". In our cohort, almost half of the cases remain undiagnosed. To further investigate those cases, we are using different approaches.

The first methodology is to periodically re-analyse negative cases: it has been widely demonstrated that re-analyses of ES data after one to two years from the first screening can even double the molecular diagnostic yield[48–50]. This increase is mainly due to the rapid improvement in knowledge of genetic causes of disease, followed by novel phenotypes associated to known genes, upgraded variant classification, improvement of bioinformatic pipelines and access to more detailed clinical information[48]. For our cohort, we "manually" check unsolved cases after 1 to 2 years from the first analysis, however we are working on a semi-automatic pipeline with the group of Dr Paolo Uva (Giannina Gaslini Hospital, Genova, Italy).

In our cohort, we have solved many cases thanks to re-analysis: an emblematic case is a male patient with the *de novo* variant c.1652A>G; p.(Asp551Gly) in *TLK2*, reported as "Case 1-Family 1" in Chapter 3 and by Pavinato *et al*[23]. This case was first analysed at the beginning of 2018 and reported as a VUS in a gene not associated with a disorder. A re-analysis of this patient one year later, showed the same variant again, this time classified as a "likely pathogenic variant" and upgraded to "pathogenic" adding the *de novo* criteria. How can we explain this? Our first analysis at the beginning of 2018, *TLK2* was not associated to a disorder and the literature was not suggesting a possible role in brain development and/or function; thus, the variant was discarded. In June 2018, *TLK2* has been associated with Mental Retardation Autosomal Dominant 57 by Reijnders *et al*.[22], and re-classified in our pipeline. This re-analysis not only solved this case, but also prompted us to search for other cases in our cohort, identifying and finally finding a molecular diagnosis for four additional affected patients from the same family (Family 2 in Chapter 3[23]).

Another strategy we are routinely using to solve undiagnosed patients is to evaluate non-random X-chromosome inactivation (XCI) in the mothers of unsolved male patients and in

unsolved affected females. The hypothesis behind this test is that an unbalanced XCI could unmask previously discarded genetic variants on the X-chromosome connected both to XCI and NDD. A pilot study using this approach in our laboratory reported solving rate of 25% of analysed negative cases, thanks to the re-evaluation of X-linked variants[51]. A similar study reported results in line with ours[52].

Recently, we began also to evaluate the impact of digenic variants using the DIVAs platform from enGenome (www.engenome.com): this tool prioritizes for pathogenic combinations of variants in disease-associated genes using artificial intelligence. We started our analysis first on nine undiagnosed families with at least two affected siblings, with the idea of finding shared pathogenic digenic combinations. Our analysis is still ongoing and needs further studies, but we are confidently in reporting the identification of a potential digenic combination in two affected monozygotic twins. The two affected probands show global developmental delay, premature birth, seizures, and electroencephalography anomalies. Using DIVAs pipeline, we identified a combination of a paternal variant in the gene *QRSL1* (NM_018292.5) [c.45delA; p.(Gly16fs) and a maternal one in the gene *GATB* (NM_004564.3) [c.29G>C; p.(Cys10Ser)]. Intriguingly, both genes are associated with autosomal recessive forms of combined oxidative phosphorylation deficiency (OMIM # 618835 and # 618838) and the encoded proteins, GatA and GatB, work together in the GatCAB complex[53]. To date, a compound heterozygous mutation in *GATB* has been reported only in two siblings, resulting in death in the first days of life[53], while compound heterozygous or homozygous variants in *QRSL1* have been reported in 5 families as associated with a lethal disease[53–55] and recently in one patient with a neurodevelopmental disorder with syndromic features[56]. Our hypothesis is that we are facing a digenic disorder, with clinical features less severe than the ones previously reported as associated with variants in single genes. We are still in a primordial stage of the study, but we confirmed the absence of the variant combination in the healthy sister, and we are currently waiting for laboratory and biochemical studies on patients' blood. The next steps of our study will include experiments on human induced pluripotent stem cells (hiPSCs) derived from patients' peripheral blood mononuclear cell (PBMC) and on overexpressing cellular models.

Finally, further strategies have been planned in the projects "Multiomic strategies to implement the diagnostic workflow of rare diseases" and "Unveiling the hidden side of NEUrodevelopmental DIsorder Genetics (NEUDIG)" which just began in our group. These works will integrate different *-omics* technologies, including short-read genome sequencing, transcriptomics, and methylation analysis, together with functional analyses on patient-derived

hiPSCs. Our final goal will be to increase the diagnostic yield, identifying peculiar and/or exclusive mechanisms of action as well as novel disease genes.

Chapter 2: Novel genes involved in neurodevelopmental disorders

This section will focus on novel NDD-related genes studied in-depth during my PhD, as principal investigator (*CAPRIN1*, published in 2022[21]) or co-principal investigator (*RPH3A,* under review). I have also collaborated to the identification of the novel NDD-genes *GLRA2*[42], *RICTOR*, *RNF40*, *GPRASP1*, *SF3B3* and *ALG5* (ongoing studies, not reported in this thesis).

2.1 CAPRIN1, a novel gene involved in language impairment, ADHD and ASD

The Cell cycle associated protein 1 (*CAPRIN1*; MIM*601178) gene encodes an RNA-binding protein (RBP) that is ubiquitous and is highly expressed in the brain (Fig. S2A). While it is mainly known for its role in cancer[57], evidence implicating CAPRIN1 in brain development and function are emerging from animal studies. Homozygous *Caprin1* knock-out (ko) mice (*Caprin1*$^{-/-}$) are not viable due to respiratory failure after birth[58]. *Caprin1* haploinsufficient mice (*Caprin*$^{+/-}$) show reduced sociability, attenuated preference for social novelty or novel objects, and deficits in reversal learning compared to controls[59]. Further, *Caprin1* deficiency in neurons causes a reduction of the complexity of the dendritic arborization and smaller dendritic spines, leading to reduced postsynaptic response to stimulation, and impaired long-term memory[58,60]. Molecularly, *Caprin1* deficiency in neurons causes altered localization of a subset of mRNAs encoding proteins fundamental for synaptic plasticity, including AMPA receptors[4,5].

CAPRIN1 interacts with two other RNA-binding proteins critical for the formation of ribonucleoparticles (RNP) known as RNA granules: FMRP[61] and G3BP1[62,63]. Moreover, CAPRIN1 is suggested to be a scaffold protein able to mediate the formation of distinct RNP complexes[62]. CAPRIN1 localizes in neuronal RNA granules at dendrites[60,64], where it binds key mRNAs involved in neuronal activity and synaptic plasticity, such as CamKIIα, BDNF, CREB, MAP2 and TrkB[64]. Following synaptic stimulations, CAPRIN1 can repress mRNA local translation through its N-terminal domain[64]. Nevertheless, the overexpression of CAPRIN1 results in a global inhibition in protein synthesis, through the phosphorylation of eukaryotic translation initiation factor 2alpha (eIF-2α)[65].

Previous exome-wide screenings in neurodevelopmental disorders have identified a few cases with nonsense variants in *CAPRIN1* [11,66], suggesting that the gene might be a candidate risk gene for ASD. Also, in a recent integrated analysis on over 63,000 exomes (over 20,000 with ASD)[67], *CAPRIN1* was identified as one of the 183 genes associated with ASD, with a false discovery rate (FDR) ≤ 0.05 (*CAPRIN1* TADA-ASD FDR=0.01). Interestingly, the association of the gene with a broader spectrum of neurodevelopmental disorders (NDDs) was even stronger (*CAPRIN1* TADA-

NDD FDR=0.00065) and *CAPRIN1* was listed among the 373 genes associated with general NDDs[67].

Here, we integrate human genetic data and functional evidence on patient-derived cells and human neuronal models to describe a novel autosomal dominant neurodevelopmental disorder associated with *CAPRIN1* haploinsufficiency.

Materials and methods

Exome and genome sequencing, variant calling and prioritization

We identified *CAPRIN1* variants in clinical diagnostic or research settings by exome sequencing (ES), genome sequencing (GS) or array-CGH (60 K; Agilent Technologies, Santa Clara, CA, USA) (see Supplementary material and methods). ES or GS trio raw data were processed and analysed using in-house implemented pipelines. High-quality variants were filtered against public databases (dbSNP150 and gnomAD ver.2.1.1) so that only variants with unknown frequency or having MAF <0.1% were considered. A further variant stratification in conformity with the American College of Medical Genetics and Genomics (ACMG) guideline[34] , also considering mode of inheritance and functional *in silico* prediction of impact, allowed us considering the final set of variants for possible associations with the phenotype. All the variants are referred to genome assembly GRCh37 (hg19), NM_005898.5. Informed consent was obtained from the participating families and the study protocol was approved by each institutes internal Ethics Committees, according to the Declaration of Helsinki.

Cell cultures

PBMCs were isolated from whole blood using Histopaque-1077 (Sigma-Aldrich, St. Louis, MS, USA) and subsequently immortalized with EBV. Resultant lymphoblastoid cell line (LCL) was cultured in RPMI medium (GIBCO, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 1% Pen-Strep and 1% L-Glutamine. Fibroblasts were obtained from a skin biopsy and cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, 1% Pen-Strep and 1 mM sodium pyruvate (Thermo Fisher Scientific) at 37°C, 5% CO_2.

DNA and RNA extraction and analysis

DNA extraction from cells was performed by incubating cell pellets in buffer supplemented with proteinase K (DNeasy Blood & Tissue Kit; Qiagen, Hilden, Germany) for 90 min at 60°C, followed by 10 min at 95°C. After centrifugation at 14000 rpm for 10 min, supernatant was recovered and incubated for 30 min at 37°C and subsequently quantified at NanoDrop (Thermo Fisher Scientific).

Total RNA was extracted from fibroblasts, PBMCs or LCL using the Direct-Zol RNA MiniPrep system (Zymo Research, Irvine, CA, USA); genomic DNA was removed by treatment with DNase I (Sigma-Aldrich), following the manufacturer's protocol. To evaluate gene expression in different tissues, commercially available total RNA from fetal total brain and adult total cerebellum, brain, frontal cortex, placenta, kidney, stomach, heart, and lung were used (Applied Biosystems, Thermo Fisher Scientific).

Complementary DNA (cDNA) was generated using the M-MLV Reverse Transcriptase kit (Invitrogen, Thermo Fisher Scientific). Expression of *CAPRIN1* was measured using the Universal Probe Library system (probe #67, Roche Diagnostics, Basel, Switzerland) and two specific intron-spanning primers. Reference genes *TBP* and *GUSB* were measured with human predesigned assays (Applied Biosystems, Thermo Fisher Scientific). Reactions were carried out in triplicate on an ABI 7500 real-time PCR machine using the ABI 2X TaqMan Universal PCR Master Mix II, according to the manufacturer's instructions (Thermo Fisher Scientific).

Analysis of splicing variants

Total RNA was extracted from LCLs, and reverse transcribed as described above; cDNA was amplified with touchdown PCR using KAPA Taq PCR Kit (Roche Diagnostics, Basel Switzerland). Amplimers were visualized into 2% TBE-agarose gel using ChemiDoc Imaging System (BioRad, Hercules, CA, USA). Bands were gel-excised, and the DNA extracted by GenElute Gel Extraction Kit (Sigma-Aldrich). Fragments were sequenced by Sanger method.

Evaluation of nonsense-mediated decay (NMD)

9×10^4 fibroblasts from cases #8 and #9 were seeded in 6-well plate and cultured until 80% of confluence. From case #3, 6×10^5 lymphoblastoid cells were used. Cells were then treated with 100 µg/ml cycloheximide for 6 hours. After washing with 1x Phosphate Buffered Saline (PBS), cells were detached with 1x Trypsin (Gibco) and prepared for RNA extraction. cDNAs were sequenced by Sanger method.

CRISPR/Cas9 heterozygous *CAPRIN1* knockout generation

Human induced pluripotent stem cells (hiPSCs) derived from HUVEC[68] were used to generate an isogenic hiPSC line carrying a heterozygous *CAPRIN1* deletion based on a published protocol[69]. Briefly, HUVEC hiPSCs were cultured in mTeSR Plus medium (StemCell Technologies, Vancouver, Canada) supplemented with 10 µM Y-27632 (Selleckchem, Houston, TX, USA) at a confluence of 2×10^4 cells/cm^2. After 48 hours cells were transfected with FuGENE HD Transfection Reagent (Promega, Madison, WI, USA) with a Cas9 expression vector (pX330-hCas9-long-chimeric-grna-g 2p) in which a *CAPRIN1*-targeting gRNA (5'-ACCGTTTGCATGGATTGGAA-3') was subcloned. Starting 24 hours post-transfection, the cells were selected for 48 hours with 1 µg/ml Puromycin

(ThermoFisher Scientific) and allowed to recover for 24 hours, before being split with Accutase (ThermoFisher Scientific) and seeded in mTeSR Plus medium with 10 µM Y-27632 into a 96 wells plate at single cell density. After 12 days, colonies were split into 24 wells plates. When ready for passaging, cells were divided into two aliquots: one was frozen and the other used for DNA extraction with QuickExtract DNA Extraction Solution (Lucigen, Middleton, WI, USA). A 326-bps region targeted by the gRNA was amplified by PCR (CAPRIN1-SA_F1: 5'-TCCCTTCCTGCTGCGTCTCA-3'; CAPRIN1-SA_R1: 5'-ACAGGCACATCACTCATAAGGCTAA-3') and genome editing was screened by T7 Endonuclease (NEB) assay. In case of positive result, a 901-bps region was sequenced by Sanger sequencing (CAPRIN1-E14_F1: 5'-TCATCCACAAGTGAGGGGTACACA-3'; CAPRIN1-E14_R1: 5'-AGGTGCTTGCTCAAGAATCAAGGTA-3'). Clones harbouring heterozygous loss-of-function variants were expanded for further studies and are referred as *CAPRIN1*$^{+/-}$, while wild-type cells are indicated as CTRL.

hiPSCs differentiation into cortical neurons

hiPSCs were differentiated into cortical neurons as previously described[70,71] with some modifications. Briefly, hiPSCs were grown as usual until confluence. A chess pattern was created using a needle, to have squares of 2-3 mm/each; subsequently they were washed with PBS and left in collagenase [3 mg/ml] for ~30 min. Detached cells were then gently resuspended into Neuronal Induction Medium #1 (see Supplementary) and kept in a 5% CO_2 incubator with 35 rpm shaking for 10 days. Medium was changed every other day. At day 10, embryoid bodies were plated on poly-Ornhitin/Laminin coated plates in Neuronal Induction Medium #2 (see Supplementary) and medium was changed every other day until confluence. Then, cells were detached and frozen or plated for neuronal differentiation in PEI/Laminin coated plates with Neuronal Differentiation Medium (see details in supplementary materials and methods). Cells were fed every three days changing half of the medium (for detailed description see Supplementary Material and Methods)

Western blot

Total proteins were extracted from cellular pellet with RIPA buffer (50 mM Tris-HCl pH 7.5; 150 mM NaCl; 1% NP-40; 0.5% Sodium Deoxycholate) supplemented with DTT 0.1 M, EDTA 0.5 M and 100x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). Ten micrograms of proteins were diluted into 4X LDS sample buffer and 10X Sample Reducing Agent and were electrophoresed on 4–12% Bis-Tris Protein Gels (Thermo Fisher Scientific).

Nitrocellulose membranes 0.45 µm were stained with MemCode Reversible Protein Stain Kit (Thermo Fisher Scientific) and immunoblotted. Antibodies are listed in the Supplementary Material and Methods. Images were acquired with ChemiDoc Imaging System and analysed with

ImageLab software (BioRad), using volume tools quantification with global subtraction method. CAPRIN1 protein level was normalized on Vinculin (MemCode normalization gave similar results).

Immunofluorescence and imaging

When not otherwise stated, immunofluorescence of hiPSCs-derived neurons was performed with our standard protocol. Briefly, neurons were plated on coverslips at low confluence (50,000 cells/cm^2). The fixation was performed with 4% PFA 4% sucrose at room temperature for 15 minutes, followed by three washes 5 minutes/each with PBS. Coverslips were then stored at 4°C until immunostaining or immediately processed. Cells were permeabilized using PBS 0.1% Triton for 5 minutes and blocked with PBS 5% BSA for 30 minutes at room temperature. Primary antibodies were diluted in PBS 1% BSA and left overnight at 4°C; secondary antibodies were diluted in PBS 1% BSA and left 2 hours at room temperature protected from light. A full list of antibodies is given in Supplementary Material and Methods. For some experiments, cells were subsequently stained with Phalloidin for 45 minutes, prior to being mounted with Fluoromount G with DAPI (Invitrogen). When not otherwise stated, images were acquired using Zeiss Imager M2 microscope, with or without Zeiss Apotome grid, and analysed using Fiji/ImageJ.

Calcium recording

Neuronal progenitors from 3 independent differentiations were plated on coverslips in a 6 well plate (50,000 cells/cm^2), in 2 ml of Neuronal Differentiation Medium and they were cultured for 10 or 14 days. The day of the recording, coverslips were moved to microscope chambers and 0.5µl F-127 (0.2%) and 0.5µl of Fluo-4 (Invitrogen) [2 µM] were added in 250 µl of pre-conditioned medium. After 25 min of incubation (5% CO_2, 37°C), the medium was removed, cells were washed with PBS and 500 µl of pre-conditioned medium without F-127 and Fluo-4 were added. Coverslips were immediately analysed using a Leica Thunder Microscope (5% CO_2, 37°C), 20x plan magnification. Each analysed segment of the coverslip was recorded for 2 min using *live* imaging. At least 446 cells/genotype and three recordings were analysed using Fiji/ImageJ and summing the IntDen signal of each frame normalized on the lowest IntDen (F_0) of each cell, as previously described[72].

Reactive oxygen species analysis

Reactive oxygen species (ROS) were detected using the fluorogenic probe CellROX Green Reagent (Thermo Fisher Scientific). Eighteen days after starting the neuronal differentiation (cells plated on coverslips, 50,000 cells/cm^2), 0.5 µl of reagent were added to 250 µl of pre-conditioned medium. After 30 minutes of incubation (5% CO_2, 37°C), medium was removed and cells were washed three times with PBS, prior to being fixed with 4% PFA 4% sucrose for 15

minutes at room temperature. Cells were then washed three times with PBS and once with water and mounted with Fluoromount G with DAPI (Invitrogen); 20x images were acquired using Zeiss Imager M2 microscope, within 24 hours from experiment. Analysis was performed using Fiji/ImageJ, setting ROIs on the DAPI signal (channel 1) and measuring IntDen on the CellROX signal (channel 2); signals were normalized on the average IntDen of backgrounds. Twenty-one images/genotype were analysed, for a total of a minimum of 697 cells/genotype.

SUrface SEnsing of Translation (SUnSET assay)

Protein synthesis was measured by SUnSET assay[73]. Differentiated neurons (post-ND day 4 and 14) were incubated with 1 µM puromycin for 30 min. Puromycin labeled peptides were detected by western blotting or immunostaining using an anti-Puromycin antibody. For western blotting, Puromycin intensity was normalized on β-Actin and on stain free gel activation signal (ChemiDoc Imaging System, 5 minutes, faint bands), the latter performed after the electrophoresis and before the transfer. Signal intensities were analyzed using Bio-Rad ImageLab software, quantifying the whole lane for anti-puromycin and stain free gel activation signal and the single bands for anti-β-Actin (Fig. S8B). Similar results were obtained for Puromycin normalized on stain free gel activation signal and on β-Actin signal.

For immunostaining, images were captured with Zeiss Imager M2 and analyzed using Fiji/ImageJ. Lookup table (LUT) "royal" was used to better visualize signal intensities; Puromycin gray values were registered along the dendrite of 6 different neurons, from 0 to 50 µm of distance from the soma. Bins of 10 µm/each -starting from 5 µm of distance from the soma- where then generated and the mean intensity of each bin was used for statistical analysis.

Micro-electrode array (MEA)

Recordings were performed using a multi-electrode array (Multi Channel Systems, Reutlingen, Germany) at different time points, using a 24-well plate (12 gold electrodes on epoxy material per well). 75,000 cells/well were plated, on a PEI/Laminin coating. Measurements were done before the medium was changed. Data acquisition and analysis was performed through the Multi-Channel Suite software (Multi Channel Systems), setting the threshold for spike detection at ±10 µV, a band-pass filter with 100 Hz and 3 kHz cut-off frequencies. Spike detection was performed using an adaptive threshold set to 5.5 times the standard deviation of the estimated noise on each electrode. Each plate first rested for 2 minutes in the recorder. Experiments were performed in a non-humidified incubator at 37°C, recording for 2 minutes. Neurons from three independent differentiation batches were measured (CTRL N=14 wells (two biological replicates, five technical replicates/each and one biological replicate with four technical replicate); *CAPRIN1*$^{+/-}$ N= 9 wells (three biological replicates, three technical replicates/each)).

Statistical analysis

For each experiment, biological triplicate with at least two technical replicates were performed. Except where otherwise indicated, data were analysed with Prism-GraphPad Software performing Student t-test with Welch's correction; p-value was indicated as follows: not significative (ns)= $p > 0.05$; *= $p \leq 0.05$; ** $p \leq 0.01$; ***= $p \leq 0.001$; ****= $p \leq 0.0001$.

In silico protein structure modelling

CAPRIN 1 protein was modelled using Protein Homology/analogY Recognition Engine (PHYRE) version 2[74], choosing intensive modelling mode option. A pdb file was generated for both wild-type and mutant protein carrying the c.279+1G>T variant and imported in EzMol[75] interface server (ver.2.1); the colour code was generated from EzMol, and it is based on the isoelectric point (red: high/positive, white: neutral/zero; blue: low/negative). The resulting image was used to visualize the 3D protein conformation. ClusPro[76–79] protein-protein docking server was used to analyse dimer formation.

Methylation data generation and analysis

Peripheral blood DNA was extracted using standard techniques. Bisulfite conversion was performed with 500 ng of genomic DNA using the Zymo EZ-96 DNA Methylation Kit (D5004), and bisulfite-converted DNA was used as input to the Illumina Infinium HumanMethylation450 (450K array) or MethylationEPIC BeadChip array (EPIC array). Array data were generated according to the manufacturer's protocol. Sample quality control was performed using the R minfi package version 1.35.2.34. The data analysis pipeline was performed as previously described[39].

Results

CAPRIN1 variants and patients' clinics

A *de novo* ~1.4 Mb deletion encompassing *CAPRIN1* was detected in a routine diagnostic array CGH screening for patients with neurodevelopmental disorders (Fig. 1A). The deletion spans eight genes, including *CAPRIN1*. *CAPRIN1* was the only one predicted to be haploinsufficient (gnomAD ver2.1.1, pLI score=0.97)[80] and had prior evidence of variants in individuals diagnosed with neurodevelopmental disorders[11,66,67]. Using GeneMatcher[40] and MSSNG Databases, we identified by exome or genome sequencing eleven additional cases with nonsense or splicing variants affecting donor/acceptor sites in *CAPRIN1* (Fig. 1A, Table 1, Fig. S1). A nonsense variant identified in case #9 was paternally inherited and the father was also affected (case #10). All other variants but one (case #2) were *de novo*, absent in the control population (GnomAD ver 2.1.1), and classified as pathogenic using the ACMG guidelines[34,35]. Two additional cases with

loss-of-function (LoF) variants in *CAPRIN1* were already reported in the literature in large genetic screenings of patients diagnosed with ASD[11,66] (Fig. 1A).

Age of cases ranged from 3 to 48 years old, and all received a clinical diagnosis of neurodevelopmental disorder. Six were born at term with normal parameters after uneventful pregnancies and delivery. Respiratory concerns were reported in six cases, and included foetal respiratory distress, postnatal saturation dips, obstructive sleep apnoea, sudden chocking spells and asthma (Table 1 and S1).

An overall representation of *CAPRIN1*-associated phenotypes is shown in Fig. 1B. Language impairment was observed in all cases, with most cases speaking very few words at last evaluations. Some of the affected individuals were supported with specific tools, as case #1, who was helped using a personal picture agenda (Video S1). Stereotyped movements and behaviours were also observed in some of the patients (Video S2, case #1).

Intellectual and/or learning disabilities from mild to very severe were reported for all cases but case #1 and #6; ASD was diagnosed in eight patients (67%). Nine patients out of eleven (82%) met diagnostic criteria for attention deficit hyperactivity disorder (ADHD). Four patients had epileptic seizures with different severity: cases #2 and #4 showed absence type seizures, resolved respectively after oxcarbazepine and ADHD treatment. Case #11 had infantile spams beginning at 9 months of age, which became absence epilepsy (Video S3) at last evaluation and was treated with sodium valproate and carbamazepine. Patient #12 suffered from secondary generalized epileptic seizures during sleep, treated with valproate with discrete control. Six patients also showed mild skeletal limb malformations, with fifth digit clinodactyly as prevalent feature (Fig. S2B); two patients had scoliosis and mild kyphoscoliosis. Brain magnetic resonance imaging (MRI) was available for six patients, and all the identified anomalies were considered without clinical significance (Fig. S2C). Gastrointestinal issues (2/12), feeding difficulties (2/12) and ocular anomalies (4/12) were reported; no facial dysmorphism was evident, although some features, such as broad nasal tip, low hanging columella and thin lip recurred in different cases (Fig. 1C and S2B, Table S2). Notably, mild hearing loss has been observed in cases #3 and #10. This finding supports the recent evidence that lack of *Caprin1* in inner ear is related to progressive hearing loss and reduced recovery from noise exposure in mice[81]. A summary of the prevalent features is reported in Table 1 and S1, and a detailed clinical description is available in Supplementary Data.

CAPRIN1 expression and splicing analysis

To verify the consequences of the variants on *CAPRIN1* expression, we measured *CAPRIN1* mRNA levels by qRT-PCR in fibroblasts and PBMCs from patients, and found that it was reduced, compatible with the loss of one functional allele (Fig. 2A). In cases #8 and #12, we detected the exclusive expression of the reference allele by Sanger sequencing. To investigate the occurrence of nonsense mediated decay (NMD) for premature stop codons, we treated fibroblasts with cycloheximide (CHX) and showed a partial rescue in the expression of the mutant allele by Sanger sequencing (Fig. 2B). We confirmed a significant reduction of CAPRIN1 protein in fibroblasts derived from patients by western blot (Fig. 2C).

In cases #2 and #3, the c.279+1G>T and c.879G>A variants were predicted to alter the donor splice sites, since they are at the first nucleotide of intron 3 and at the last base of exon 8[82]. Analysis of the cDNA showed skipping of exon 3 for c.279+1G>T and of exon 8 for c.879G>A (Fig. 2D). In the first case, the mature mRNA is predicted to encode a protein with an in-frame deletion of 31 aa (63 bp) occurring in a highly conserved region. Although the specific function of the encoded region is unknown (amino acids 73-93), it falls within the HR1 domain of CAPRIN1 (amino acids 56-248), which falls within the homodimerization domain (amino acids 132-251). The large concave negatively charged surface formed is a protein-protein binding pocket and it mediates the interaction with up to two molecules of FMRP[62]. *In silico* analysis of mutant protein (CAPRIN1 lacking exon 3) showed an alteration in both its tridimensional structure and its ability to homodimerize (Fig. S3B). This could reduce protein stability and thus lead to the reduced abundance that we observed (Fig. S3A).

The second variant causes the loss of 53 nucleotides, resulting in an out-of-frame protein. Treatment of patient cells with cycloheximide (CHX) restored expression of the mutant allele that was absent without treatment, suggesting that it undergoes NMD (Fig. 2D, right panel). Reduced CAPRIN1 protein levels were also observed (Fig. S3A).

Overall, CAPRIN1 protein levels were clearly reduced in all samples analysed, supporting haploinsufficiency as disease mechanism. This finding is in line with the type of mutations found in the cases and the high probability of intolerance to heterozygous LoF variation (pLI) from the Genome Aggregation Database (gnomAD; https://gnomad.broadinstitute.org)[83].

CAPRIN1 loss impairs neuronal structure organization and processes length

To analyse the impact of *CAPRIN 1* haploinsufficiency in human neuronal cells, we generated an isogenic *CAPRIN1*$^{+/-}$ hiPSC line using the wild-type HUVEC hiPSC line and CRISPR/Cas9 technology[68] (Fig. S4A and B). *CAPRIN1*$^{+/-}$ hiPSCs were differentiated into cortical neurons, using established protocols[70,84] with some modifications (Fig. S4C). CAPRIN1 was well expressed in

neuronal rosettes and in late cortical progenitors, and no significant differences in size were observed in *CAPRIN1*$^{+/-}$ neuronal rosettes compared to control (Fig. 3A and Fig. S5A). No morphological alterations were observed in neural progenitor cells in further passages either (Fig. S5B). In control neurons, CAPRIN1 localizes predominantly in the soma in immature stages and is then increasingly expressed in dendrites during development (Fig. S6). *CAPRIN1*$^{+/-}$ immature neurons showed an alteration of neuronal structure, accompanied by reduced neurites length (Fig. 3B and C). This is consistent with previous observations, suggesting a reduction in dendrite length and neuronal connectivity in murine *Caprin1*$^{-/-}$ primary embryonic neurons[58].

Increased neuronal degeneration, calcium signal and reactive oxygen species in *CAPRIN1*$^{+/-}$ neurons

During time-course analysis of neuronal differentiation, we observed a progressive degeneration of neuronal processes in *CAPRIN1*$^{+/-}$ neurons, which began with the formation of small clusters of shrinking neuronal cells[85] between days 17 and 19 post neuronal differentiation (ND) in 5 independent cultures (Fig. 3D and S7A). To understand the mechanisms behind this cellular phenotype, we investigated the possibility of calcium signal alterations. Impaired Ca^{2+} homeostasis and overload have been described as triggering or contributing to neuronal cell death[86–88]. Using a fluorescent tracer, we recorded Ca^{2+} signals in neurons 10- and 14-days post-ND, using *live* imaging (Video S4). We chose these time points prior to cell death to minimize confounding factors and to study events leading to neuronal degeneration. Recordings for 2 minutes under controlled CO_2 and temperature conditions showed an increase in Ca^{2+} signals in *CAPRIN1*$^{+/-}$ compared to control neurons, both at 10- and 14-days post-ND (Fig. 4A).

Neuronal death has also been attributed to the induction of reactive oxygen species (ROS)[89]. Thus, we tested whether the elevated Ca^{2+} levels could contribute to ROS induction[90,91]. We analysed neurons 18 days post-ND, to capture the events immediately prior to neuronal death. Cells were treated with CellROX™ for 30 min and then fixed and counterstained with DAPI for immunofluorescence analysis. Notably, we observed a significant increase in CellROX signal in *CAPRIN1*$^{+/-}$ cells compared to controls (Fig. 4B), confirming an increase in ROS after Ca^{2+} overload.

Loss of *CAPRIN1* causes an increased global mRNA translation in neurons

CAPRIN1 has been suggested to be a translational inhibitor, and its overexpression results in global reduction of protein synthesis in HeLa cells[92]. Interestingly, CAPRIN1 interacts with other

proteins involved in neurodevelopment, including Fragile-X Mental Retardation Protein (FMRP)[62] and the Cytoplasmic FMRP-Interacting Protein 1 (CYFIP1)[93], both of which are involved in the regulation of mRNA translation and formation of dendritic spines[93,94].

Protein synthesis is tightly regulated in neurons, with spatial and temporal regulation of gene expression[95]. To understand whether a lack of CAPRIN1 causes an increase of protein synthesis, we performed the surface-sensing-of-translation (SUnSET) assay. This method tests global protein synthesis based on puromycin incorporation and is commonly used to measure relative rates of *de novo* protein synthesis[73]. Treatment of hiPSCs-derived neurons with 1 µg/ml puromycin for 30 minutes resulted in a slightly increased global translation in *CAPRIN1*$^{+/-}$ neurons 4 days post-ND (p= 0.228) (Fig. S8A), which became significantly higher 14 days post-ND (p= 0.03) (Fig. 5A). Interestingly, a difference in the distribution of the puromycin signal along the neurons was also observed (Fig. 5B-C), both 4 and 10 days post-ND. In control neurons, the signal was higher near the soma and then slowly decreased along the dendrite; similarly, *CAPRIN1*$^{+/-}$ cells showed higher puromycin fluorescence intensity the cell body, which decreased rapidly in the dendrite starting at ~5 µm from the soma, much more intense than in controls. These data suggest that *CAPRIN1* haploinsufficiency might also impact local protein synthesis.

CAPRIN1$^{+/-}$ neurons display abnormal spontaneous firing properties

In *Caprin1*$^{+/-}$ mice, neurons had a reduced firing activity[59]. To investigate whether the impaired neuronal structure organization affects network activity during neuronal maturation also in human *CAPRIN1*$^{+/-}$ neurons, we compared the electrical properties of *CAPRIN1*$^{+/-}$ hiPSCs-derived neurons and controls using micro-electrode arrays (MEAs)[70,71]. Spontaneous firing of neurons was monitored from day 0 to day 14 post-ND to assess stages of network maturation (Fig. 6A). Both control and *CAPRIN*$^{+/-}$ neurons showed firing properties already at 4 days post-ND. The spike rate and spike count were significantly lower in *CAPRIN1*$^{+/-}$ neurons versus controls, both at immature (post-ND days 4 and 7) and more mature stages (post-ND days 10 and 14) (Fig. 6B and C). In addition, the burst count 10- and 14-days post-ND, was significantly lower in *CAPRIN1*$^{+/-}$ neurons, with the absence of bursts in most of the wells. Interestingly, also the duration of the detected bursts was shorter in *CAPRIN1*$^{+/-}$ neurons compared to controls (Fig. 6D and E).

Overall, these results confirmed a reduced spontaneous activity (spike rate and count) and decreased synchronization of *CAPRIN1*$^{+/-}$ neurons (burst count).

CAPRIN1 haploinsufficiency is associated with a detectable episignature in patients' DNA

In the past years, it has been widely demonstrated that a possible functional consequence of pathogenic variants in patients with Mendelian disorders, including NDDs, is the alteration of genomic DNA methylation. Indeed, changes in DNA methylation patterns -indicated as episignatures-, are emerging as highly accurate and stable biomarkers in a growing number of Mendelian disorders[38,39,96]. To evaluate the possibility of identifying an episignature in patients with LoF variants in *CAPRIN1,* we took advantage of an ongoing collaboration with the EpiSign consortium (Dr Bekim Sadikovic). The analysis of 5 patients in our cohort was sufficient to generate preliminary supporting data about the presence of a recognizable methylation profile; the addition of 7 more patients to the cohort (Table 2) confirmed the preliminary data, showing that of training data shows all *CAPRIN1* samples grouped separately from controls (Fig. 7A-D). Training of samples against controls (healthy cases and cases affected by other NDDs not associated with an episignature) showed high sensitivity of the analysis. Furthermore, training of *CAPRIN1* samples against methylation profiles derived from cases with other NDDs previously associated with an identifiable episignature, showed high specificity of the model, showing a well detectable separation of *CAPRIN1* samples from the other ones (Fig. 7E-F). Analysis of further seven cases, including patients with missense variants, is ongoing.

Discussion

CAPRIN1 is a candidate gene for ASD (high confidence SFARI 1; https://gene.sfari.org/): a *de novo* nonsense variant was identified in a male ASD proband[66] and more recently, a second nonsense *CAPRIN1* variant was found in a genome sequencing screening of families with multiple ASD-affected children [11]. In a recent combined analysis of 33 different ASD cohorts[67], *CAPRIN1* was identified as a risk gene associated with ASD and NDDs[67]. However, a causal relation between the gene and neurodevelopmental disorders was still missing. Our work provides the first clinical characterization of the individuals with *CAPRIN1* haploinsufficiency and supports a key role for *CAPRIN1* in human neuronal development and activity.

Our data point to a critical role of CAPRIN1 in neuronal morphogenesis and survival. *CAPRIN1*$^{+/-}$ hiPSC-derived neurons have aberrant neuronal organization, reduced length of neuronal processes, progressive neuronal degeneration, and premature cell death beginning 17-19 days post-ND. In *Caprin1*$^{-/-}$ cortical neurons, the length and number of dendrites, and the formation of neuronal networks were reduced compared to wild-type. Interestingly, network area was always smaller compared to wild-type neurons, but it degenerated during neuronal maturation, particularly after 22 DIV[58]. Similarly, in our cell model, we observed an aberrant neuronal organization and reduced length of neuronal processes (Fig. 3B and C). In the mouse model, the

lack of *Caprin1* caused increased cell death both *in vitro* - in cortical neurons from 10 days in vitro (DIV) - and *in vivo*, as demonstrated by an increased cleaved caspase-3 signals in brain slices [58].

One of the proposed mechanisms for these cellular phenotypes is decreased of Na^+/K^+ ATPase (NKA) activity[58]. Inhibition of NKA activity results in increased sensitivity to glutamate excitotoxicity, which subsequently causes an increase of intracellular Na^+ and Ca^{2+} ions[97]. In addition, inhibition of NKA can partially suppress excitatory postsynaptic potentials (EPSPs) and reduce surface concentrations of AMPA receptors in neurons[98]. In *CAPRIN1*$^{+/-}$ human neurons, we observed a Ca^{2+} overload 10- and 14-days post-ND and an overall reducing firing rate and synchronization on neuronal activity. In the mouse model, heterozygous loss of *CAPRIN1* causes reduced amplitude and slope of field EPSPs in hippocampal slices[60]. Altogether, our and published data support the hypothesis that previously reported impairment of NKA activity[58] impacts Ca^{2+} homeostasis and electrical activity in *CAPRIN1*$^{+/-}$ neurons.

Further, during mouse cortical development CAPRIN1 interacts with several mRNAs, including some metal-binding proteins as selenoprotein W and metallothionein 2, involved in protection against ROS[58]. In *CAPRIN1*$^{+/-}$ hiPSC-derived neurons, we observed a significant increase in ROS production concomitant with the neuronal cell degeneration 18 days post-ND. The link between ROS, autism spectrum disorder and ADHD has been reported in the literature, and some therapies are now focusing on antioxidant drugs[99]. Also, diagnostic biomarkers investigating redox metabolism, are the subject of study in patients with ASD and ADHD, as for instance total, oxidized and free reduced glutathione levels[100] or thiol/disulphide homeostasis[101]. It will be of interest to further study increased ROS species in *CAPRIN1*$^{+/-}$ neurons, as the method used in the present work detected a global ROS signal but did not discriminate between different ROS types[89].

Another potential CAPRIN1 target that might mediate the cellular phenotypes we observe is the Defender against apoptotic cell death (DAD1), a cell death suppressor that may act downstream of the bcl-2 protein[102]; in a microarray analysis of CAPRIN1-bound mRNAs, DAD1 has been identified as an interactor[58].

Non neuronal cells lacking *CAPRIN1* do not exhibit changes in global rate of protein synthesis[63], whereas the local expression of some mRNAs have been reported as altered in a conditional model of *Caprin1* knockout mouse[60]. In our work, we showed that *CAPRIN1* haploinsufficiency is sufficient to cause a significant increase of global translation in neuronal cells (Fig. 5A) and an impairment of its localization (Fig. 5B-C), enforcing the concept that the use of the right cellular model is important to detect cell-specific phenomena. Interestingly, it has been shown CAPRIN1

selectively targets dendritic mRNAs, and an increase in mRNAs in soma but a decrease in mRNAs in dendrites has been observed in *Caprin1* conditional ko neurons[60]. This further supports the decrease in local translation in distal dendrites in *CAPRIN1*$^{+/-}$ neurons that we identified (Fig. 5B-C). A rapid response to physiological and pathological mechanisms adjusting the protein synthesis rate is also mediated by oxidative stress[46]. Both the administration of H_2O_2 and the endogenous increase of ROS can reduce cytosolic protein synthesis in a reversible manner[103,104]. It is therefore possible, the increased oxidative stress we observed is an attempt of the neuron to balance the increased translational rate caused by *CAPRIN1* loss.

The lack of impact on global translation in a non-neuronal cell line[63] suggests the loss of *CAPRIN1* can have different outcomes depending on the cell type, explaining the predominantly neurological phenotypes we observed in our patients, despite the ubiquitous expression of *CAPRIN1* in human tissues.

In our patients, we reported ASD in 67% of cases, and this clinical phenotype is compatible with the reduced sociability and weaker preference for social novelty previously observed in *Caprin1*$^{+/-}$ mice[59]. We observed intellectual or learning disability in 83% of cases. *Caprin1*$^{+/-}$ mice have difficulties in reversal learning, but no overall memory deficits[59]. However, a *Caprin1* brain conditional knockout model show deficit in spatial and fear conditioning long-term memory, as well as reduced basal and induced synaptic transmission in hippocampal neurons[60]. One third of our patients showed seizures with variable presentation. Interestingly, seizures were sporadically noted in *Caprin1* mutant mice during and after the Morris water maze and contextual fear conditioning tests, and epileptic-like fEPSP were recorded after induction of long-term potentiation (LTP)[60]. These data suggest that the epileptic phenotype could be triggered by specific events. We also noted six cases with respiratory problems, a finding consistent with the post-natal failures in breathing in the *Caprin1*$^{-/-}$ mouse model[58].

Caprin1 has recently been associated with early onset progressive hearing loss and reduced protection from noise exposure in an inner ear conditional knockout mouse[81]. We found a clinical report of mild hearing loss in in a paediatric (case #3) and an adult case (case #10). The adult case works in an environment with high exposure to loud sounds, which is associated with increased risk for hearing loss[105], supporting the data that lack of *Caprin1* is involved in recovery from audible insults[81]. It is possible that most of the affected individuals do not show hearing impairment due to the young age, but they might develop this phenotype in the future.

Despite the ubiquitous expression of *CAPRIN1* in humans, we observed prevalently neurological features; however, some minor and variable defects at the level of skeletal, respiratory, visual, and auditory systems were identified. Noteworthy, the association between diseases and

affected tissues is not always straightforward and it is common in genetic disorders to observe the impairment of only a small portion of tissues compared to the broader expression of the causal gene[106]. This could be explained by several mechanisms, including disrupted pathways, regulatory elements, post-transcriptional mechanisms, and microenvironment, that can be cell- or tissue-specific[106]. Observations of *Caprin1* heterozygous, homozygous, or conditional mutants suggest impairment of the auditory[81] and respiratory[107] systems, but there is no evidence about skeletal or ocular involvement. Non-neurological features are observed only in a fraction of patients. We cannot therefore exclude the presence of intra- or extragenic eQTLs or SNPs that contribute to the variable presentation of those phenotypes.

Variants identified in our cases were heterozygous LoF because they introduced a premature stop codon by changing an amino acid, introducing a frameshift, or altering splicing [62].

Ten cases were *de novo*, and two familial (Fig. S1). Case #9 and his father (case #10) carried the p.(Arg458Ter) variant. Both showed language delay, ADHD, mild intellectual disability/developmental delay as children, but the father had recovered most of his language impairments by adulthood. Case #2 had the c.279+1G>T variant inherited from his mother, reported healthy, and without history of intellectual disability, speech delay, neurodevelopmental/behavioural concerns, or other medical issues; she had a college education. These cases suggest the disease has a variable expressivity during life and/or incomplete penetrance; it is also tempting to speculate a sex bias, given the healthy female carrier and the observation that 9 out of 12 affected individuals were males.

Our work focused on *CAPRIN1* LoF variants (nonsense, splicing and small copy number variants). The possibility that missense variants are also associated with CAPRIN1-NDD is still to be explored. Overall, *CAPRIN1* is not anticipated to be highly sensitive to missense variations (Z score= 1.69, GnomAD)[108]; however, some specific regions of the protein are predicted to be intolerant or highly intolerant to changes, according to MetaDome server (v1, https://stuart.radboudumc.nl/metadome)[109]. Intolerant amino acids (aa) are predominantly located between aa 58-131 and 514-583, part of the Homologous Regions 1 and 2 (HR1 and HR2) domains of CAPRIN1[62]. Indeed, a p.(Pro512Leu) recurrent missense variant in the HR2 domain has been recently found in two independent patients with a neurodegenerative syndrome with ataxia[110]. This indicates that different neurologic phenotypes are associated with *CAPRIN1*, depending on the variant and its pathophysiological significance. The future inclusion of *CAPRIN1* in diagnostic genetic panels might help identifying cases with missense variants and likely expand the clinical spectrum associated with the gene.

The detectable episignature observed in patients with *CAPRIN1* LoF variants open novel scenarios for both diagnostic and research purposes. For the diagnostic purposes, the possibility to classify the variants simply using patient-derived DNA from peripheral blood is an important readout[39]. For instance, in our cohort we included a variant that was difficult to interpret, and other functional testing were difficult due to the lack of patient-derived cells: as shown in Table 2, the case MS10097 carries the variant *CAPRIN1* c.217-6A>G; p.(?), absent from the control population and predicted to be likely pathogenic[34]. Due to the compatibility of the clinical presentation, the case was included in the methylation analysis, and we observed an overlapping with the other *CAPRIN1* cases, confirming this intronic variant acts with a LoF mechanism. A further analysis on 7 additional cases is ongoing and we are including also other groups of patients, namely two cases with missense variants and a clinical phenotype partially compatible with the NDD associated with LoF variants and the two cases with the p.(Pro512Leu) variant, described as associated with a gain-of-function (GoF) mechanism of *CAPRIN1*[110] . Our final aim will be to evaluate if it is possible to distinguish between variants acting through a LoF or GoF mechanism, and eventually distinguish among different subsets of patients (i.e., patients with LoF variants +/- seizures).

Altogether, our data finally link *CAPRIN1* haploinsufficiency with an autosomal dominant neurodevelopmental disorder, characterized by language impairment, ADHD, intellectual or learning disability and ASD, with other variably associated features including respiratory, skeletal, ocular, auditory and gastrointestinal issues. Likewise, *CAPRIN1* haploinsufficiency damaged the overall organization of hiPSC derived neurons, and mediated impairment of Ca^{2+} homeostasis, neuronal activity, ROS production and global translation rate.

Figures and tables

Table 1. Clinical features of patients with variants in *CAPRIN1*.

Case	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Sex	M	M	M	F	M	F	M	M	M	M	M	F
Age at examination (years)	7	12	6	17	10	8	7	7	12	48	3	10
Language impairment	+	+	+	+	+	+	+	+	+	+	+	+
ASD	+	+	+	+	+	-	+	+	-	-	+	-
ID/LD	-	++	+	+	++	-	+++	+	++	+	+	+++
ADHD	+	+	-	+	+	+	N.A.	+	+	+	-	+
Developmental delay	-	+	+	-	+/-	-	-	-	+	-	+	-

Other behavioural features	+	-	-	+	-	+	-	-	+	+	+	+
Seizures	-	+	-	+	-	-	-	-	-	-	+	+
EEG anomalies	N.A.	+	-	-	-	N.A.	N.A.	N.A.	N.A.	N.A.	+	+
Hands and feet malformations	-	+	+	+	+	-	-	-	+	-	-	+
Breathing problems	-	-	-	+	+	+	-	+	+	+	-	-
Ocular problems	-	-	-	+	+	-	-	-	+	+	-	-
Hearing problems	-	-	+	+	-	-	-	-	-	+	-	-

\+ Reported in the patient; for ID/LD: + mild; ++ moderate; +++ severe

\- Not reported in the patient

N.A. The information is not available

M: male; F: female; ASD: autism spectrum disorders; ID: intellectual disability; LD: learning disability; ADHD: attention deficit hyperactivity disorder.

Table 2. *CAPRIN1*-patients included in EpiSign analysis

ID	**Sex**	**Age**	**Variant (hg19, NM_005898.5)**	**HPO terms**
MS6769	M	2	del11p13	HP:0000729; HP:0000750; HP:0002463; HP:0007018
MS6770	M	8	c.279+1G>T, p.(?)	HP:0000729; HP:0000750; HP:0002463; HP:0001249; HP:0007018; HP:0001250; HP:0004209; HP:0002099
MS6771	M	7	c.879G>A, p.(Glu293=)	HP:0000729; HP:0000750; HP:0002463; HP:0012758; HP:0011918
MS6772	M	7	c.1195C>T, p.(Gln399*)	HP:0000729; HP:0000750; HP:0002463; HP:0001249; HP:0007018; HP:0025116
MS6773	M	8	c.1103delGinsTATGT, p.(Gly368Valfs*5)	HP:0000729; HP:0000750; HP:0002463; HP:0001249; HP:0007018
MS10094	M	12.3	c.1372C>T, p.(Arg458*)	HP:0012758; HP:0000750; HP:0001256; HP:0007018
MS10095	M	48	c.1372C>T, p.(Arg458*)	HP:0000750; HP:0001256; HP:0007018; HP:0000739
MS10096	F	8.5	c.366+2T>C, p.(?)	HP:0000750; HP:0001256; HP:0007018; HP:0000729; HP:0000739; HP:0001250
MS10097	M	9.8	c.217-6A>G, p.(?)	HP:0012758; HP:0001252; HP:0000750; HP:0000729; HP:0002342; HP:0001250; HP:0002086
MS10098	M	3	c.1493_1496del, p.(Ser498Asnfs*10)	HP:0002509; HP:0000750; HP:0012759; HP:0000729; HP:0001249; HP:0011097
MS10099	M	12	c.910C>T, p.(Gln304*)	HP:0000729
MS10100	F	46	c.910C>T, p.(Gln304*)	HP:0000729

Figure 1. *CAPRIN1*-associated disease: variants and clinical features

(A) Schematic representation of a large ~1.4 Mb deletion spanning the *CAPRIN1* gene (top) and eleven *CAPRIN1* (NM_005898.5) loss of function variants (bottom); ten single nucleotide variants detected in this work and two variants reported in the literature[11,66] are represented on a schematic view of the protein (created with BioRender.com). Overall, they include two splicing (c.279+1G>T; c.879G>A) and ten nonsense/frameshift changes. CAPRIN1 structure was schematized according to Wu *et al.*[62]: two highly conserved homologous regions (HR1, 56-248 amino acids and HR2, 352-685 amino acids) and a less conserved glutamate-rich (E-rich) region in between are shown. Moreover, C-terminal Arg-Gly-Gly (RGG) motifs -characteristic of some RNA-binding proteins[62]- are indicated in purple. **(B)** Main neurological (top) and non-neurological (bottom) features of *CAPRIN1*-associated disease in the twelve cases described here. **(C)** Facial features of the twelve cases do not show peculiar dysmorphisms, although broad nasal tip, low-hanging columella and thin upper lip vermilion seemed recurrent. Facial features analysis by Face2Gene software (FDNA Inc., https://www.face2gene.com/)[111] does not show a peculiar overlap. Further details are provided in Supplementary Fig. 2 and Supplementary Table 1.

Figure 2. Expression analysis of *CAPRIN1* in patient-derived cells

(A) Expression of *CAPRIN1* vs. *GAPDH* is reduced in patients' fibroblasts or PBMC, measured by qRT-PCR (case#8, wt/Gln399Ter (in figure Q399*) = 0.59±0.01, N=3; case#12, wt/Gln582Ter (in figure Q582*) = 0.41±0.01, N=3; CTRLs= 1.03±0.08, N=9; case#1, wt/del11p13= 0.34±0.03, N=9 CTRLs= 1.04±0.09, N=3; mean±SEM. p=0.0007 (Q399* vs CTRLs) and <0.0001 (Q582* vs CTRLs and del11p13 vs CTRLs)). For every experiment, three different controls were used, and the experiments were performed in triplicate; controls were healthy young adults. **(B)** Sanger sequencing of the cDNA containing the c.1195C>T p.Gln399Ter and c.1744C>T p.Gln582Ter shows that the mutant allele is likely degraded by nonsense mediated decay (NMD; cDNA panels). Indeed, treating cells with cycloheximide (CHX), an NMD blocker, prior to RNA extraction, partially rescued the expression of the mutant allele (cDNA+CHX panels). As comparison, the sequences obtained from genomic DNA (gDNA panels). **(C)** CAPRIN1 protein analysis in fibroblasts of the patients carrying the c.1195C>T p.Gln399Ter and c.1744C>T p.Gln582Ter variants. Western blot shows a protein reduction compatible with a half-dose (case #8= 0.26±0.03, N=4; case #12= 0.36±0.02, N=4; CTRLs=0.99±0.06, N=10. p= <0.0001). For every experiment, three controls were used from different healthy patients (young adults), and experiments were performed a minimum of three times. Vinculin was used as a normalizer, and MemCode stain (Thermo Fisher Scientific) to confirm the efficiency of protein transferring and

the overall protein abundancy. **(D)** Analysis of the cDNA from LCLs of the c.279+1G>T (case #2) and c.879G>A (case #3) carriers. The c.279+1G>T change causes an in frame 84 bp exon 3 skipping (left panel). The c.879G>A affects the last base of exon 8 and causes out of frame skipping of exon 8 and a degradation by NMD. Exon 8 skipping could be detected by cycloheximide (CHX) treatment (right panel). Control (CTRL) LCLs derived from a young adult sex matched healthy patient.

Figure 3. Haploinsufficiency of *CAPRIN1* causes impaired neuronal organization and increased degeneration of neuronal processes

(A) Representative images of neuronal rosettes derived from CTRL and *CAPRIN1*$^{+/-}$ cells. CAPRIN1 protein (magenta) could be easily detected in these structures. No gross morphological differences were observed between CTRL and *CAPRIN1*$^{+/-}$ rosettes. Images were acquired at 40x magnification, using Z-stack mode. The neuronal marker Tuj1 (green) is used to see the rosette morphology, and the DAPI (blue) to stain the nucleus. **(B)** *CAPRIN1*$^{+/-}$ neurons 10 days post-ND showed an impaired neuronal organization (20x magnification; blue: DAPI; magenta: CAPRIN1; green: Tuj1. Scale bar=25 µm). **(C)** Quantification of neuronal processes length on 20x brightfield images, day 7 post-ND. Length was quantified using Fiji/ImageJ tool. CTRL= 134±5.62, N= 90; *CAPRIN1*$^{+/-}$= 74.8±3.55, N= 90; mean±SEM; p= <0.0001. Each dot represents a different neuronal process; images deriving from three independent differentiations were analysed. **(D)** Representative images of the neuronal death observed starting 18 days post-ND. Images were acquired 18 days post-ND (left:4x magnification, centre: 10x magnification) and 21 days post-ND (right: 20x magnification), using a brightfield microscope.

Figure 4. Calcium overload and increased oxidative stress are observed in *CAPRIN1*$^{+/-}$ neurons

(A) Quantification of calcium signal in *CAPRIN1*$^{+/-}$ neurons 10- and 14-days post-ND. Neurons were recorded for 2 minutes, and integrated density (IntDen) signal of each cell was normalized on its lowest value during this time. The right panel shows a representative image of the lowest (F_0) and maximum (F_{max}) fluorescence intensity for each cell; the lookup table "fire" was used. On the left is visible the signal of the Fluo-4 (green), better appreciable with the image on the right (day 10 post-ND: CTRL= 127±0.25, N=197; *CAPRIN1*$^{+/-}$=131±0.37, N=197; day 14 post-ND: CTRL= 125±0.67, N=354; *CAPRIN1*$^{+/-}$=138±0.45, N=361; p<0.0001). For every genotype and time point, three different coverslips were analysed, each of them deriving from an independent differentiation. **(B)** Analysis of *CAPRIN1*$^{+/-}$ neurons (day 18 post-ND) with CellROX showed a significant increase of reactive oxygen species (CTRL= $1.73x10^7$±$1.80x10^6$, N=697; *CAPRIN1*$^{+/-}$= $5.64x10^7$±$2.96x10^6$, N=1596; p<0.0001). Twenty-two images for every genotype have been analysed, from three different coverslips, each of them deriving from an independent

differentiation. On the right a panel exemplifying the difference between CTRL and *CAPRIN1*$^{+/-}$ neurons (20x magnification; blue: DAPI; green: CellROX. Scale bar= 25 µm).

Figure 5. *CAPRIN1*$^{+/-}$ neurons show increased global translation and slight alteration of local translation

(A) Western blot analysis of SUnSET assay on neurons 14 days post-ND treated with 1 µM puromycin for 30 minutes showed a significantly higher puromycin signal compared to CTRL. Puromycin-untreated cells (mock) were used as negative control for the antibody signal (CTRL=1.0±0.13; *CAPRIN1*$^{+/-}$= 1.8±0.16; mean±SEM; p= 0.02). Each CTRLs and *CAPRIN1*$^{+/-}$ lane (and dot on graph) represents an independent neuronal differentiation (N=3). **(B)** Immunofluorescence analysis of SUnSET assay on *CAPRIN1*$^{+/-}$ neurons 4 days post-ND, treated with 1µM puromycin for 30 minutes. Representative images for CTRL and *CAPRIN1*$^{+/-}$ dendrites are reported (4 days post-ND, zoom of 40x magnification, scale bar 1 µm), using the ImageJ/Fiji lookup table "royal", to better visualize the strength of the signal[112]. Magnification of the dendrites, show the puromycin fluorescence intensity from 0 to 80 µm of distance from the cell bodies; intensity of signal is maintained along the dendrite in CTRL neurons, while it decreases visibly in *CAPRIN1*$^{+/-}$ ones. An additional zoom on fragment from 15 to 30 µm of the dendrite is provided (scale bar 1 µm), showing reduced intensity of signal in *CAPRIN1*$^{+/-}$ neurons[59]. An image of the entire neuron is provided in Fig. S8C. **(C)** Quantification of Puromycin signal, indicated as gray value, along the dendrite (N=6). Both panels (left=4 days post-ND; right=14 days post-ND) show higher signal in proximity of the soma, that progressively decrease. However, the decrease is significantly higher in *CAPRIN1*$^{+/-}$ neurons, particularly 4 days post-ND. Statistical significance was tested by two-way ANOVA followed by Bonferroni post-tests; column factor was considered. P value <0.0001; mean and error (SEM) are showed.[59]

Figure 6. Neurons deficient for *CAPRIN1* show abnormal firing properties.

(A) MEA analyses in CTRL and *CAPRIN1*$^{+/-}$ neurons at different stages of maturation (day 4, 7, 10 and 14 post-ND). Each image is representative of three different wells, recorded for 2 minutes on 12 electrodes. Box and whiskers plots (5-95 percentile), outliers are indicated as dots; mean is indicated as "+"; numbers on each bar indicate the mean. Wells that never showed an activity throughout the time course have been excluded from the analysis (total CTRL N=14 wells; total *CAPRIN1*$^{+/-}$ N=9 wells). **(B)** Quantification of MEA data, reporting the spike rate (Hz) along the 2 minutes of recording (day 4 CTRL=0.66±0.14; *CAPRIN1*$^{+/-}$= 0.22±0.09; mean±SEM; p= 0.02; day 7 CTRL=0.78±0.11; *CAPRIN1*$^{+/-}$= 0.36±0.08; p= 0.006; day 10 CTRL=1.55±0.28; *CAPRIN1*$^{+/-}$= 0.66±0.13; p= 0.01; day 14 CTRL=1.82±0.29; *CAPRIN1*$^{+/-}$= 1.03±0.20; p= 0.04). **(C)** Quantification of MEA data, reporting the spikes count along the 2 minutes of recording (day 4 CTRL=0.54±0.25;

CAPRIN1$^{+/-}$= 0.16±0.10; mean±SEM; p= 0.18; day 7 CTRL=0.50±0.22; *CAPRIN1*$^{+/-}$= 0.30±0.17; p= 0.48; day 10 CTRL=6.49±1.05; *CAPRIN1*$^{+/-}$= 1.03±0.47; p= 0.002; day 14 CTRL=7.62±0.64; *CAPRIN1*$^{+/-}$= 3.30±0.98; p= 0.002). **(D)** Burst count along the 2 minutes of recording (day 4 CTRL=78.7±17.3; *CAPRIN1*$^{+/-}$= 26.7±10.9; mean±SEM; p= 0.02; day 7 CTRL=93.5±13.6; *CAPRIN1*$^{+/-}$ = 42.7±9.06; p= 0.006; day 10 CTRL=186±33.3; *CAPRIN1*$^{+/-}$= 78.8±15.7; p= 0.01; day 14 CTRL=215±35.3; *CAPRIN1*$^{+/-}$= 124±24.3; p= 0.05). **(E)** Burst duration (ms) along the 2 minutes of recording (day 7 CTRL=233±36.7; *CAPRIN1*$^{+/-}$= 186±25.3; day 10 CTRL=212±8.76; *CAPRIN1*$^{+/-}$= 223±9.51; day 14 CTRL=183±18.5; *CAPRIN1*$^{+/-}$= 147±4.49). The number of bursts in CTRL was higher compared to *CAPRIN1*$^{+/-}$, therefore the number of considered events (N) is reported. **(F)** Representative images of CTRL and *CAPRIN1*$^{+/-}$ detected spikes.

Figure 7. *CAPRIN1* haploinsufficiency is associated with an identifiable episignature in patients' DNA.

(A-B) Differentially methylated probes were selected. **(A)** p-value distribution and (**B)** volcano plot of the mean methylation differences and p-values are shown, where selected probes are in red. (**C)** Hierarchical clustering of training data shows all *CAPRIN1* samples grouping separately from controls. (**D)** MDS plot also shows clear separation of cases from controls. **(E-F)** A support vector machine model was used to classify samples and calculate probability scores. **(E)** Samples were trained against controls. **(F)** Samples were trained against controls plus samples from other neurodevelopmental disorders (NDD): 75% of controls and NDD samples were used for training (blue), and 25% for testing (grey).

Figure 1

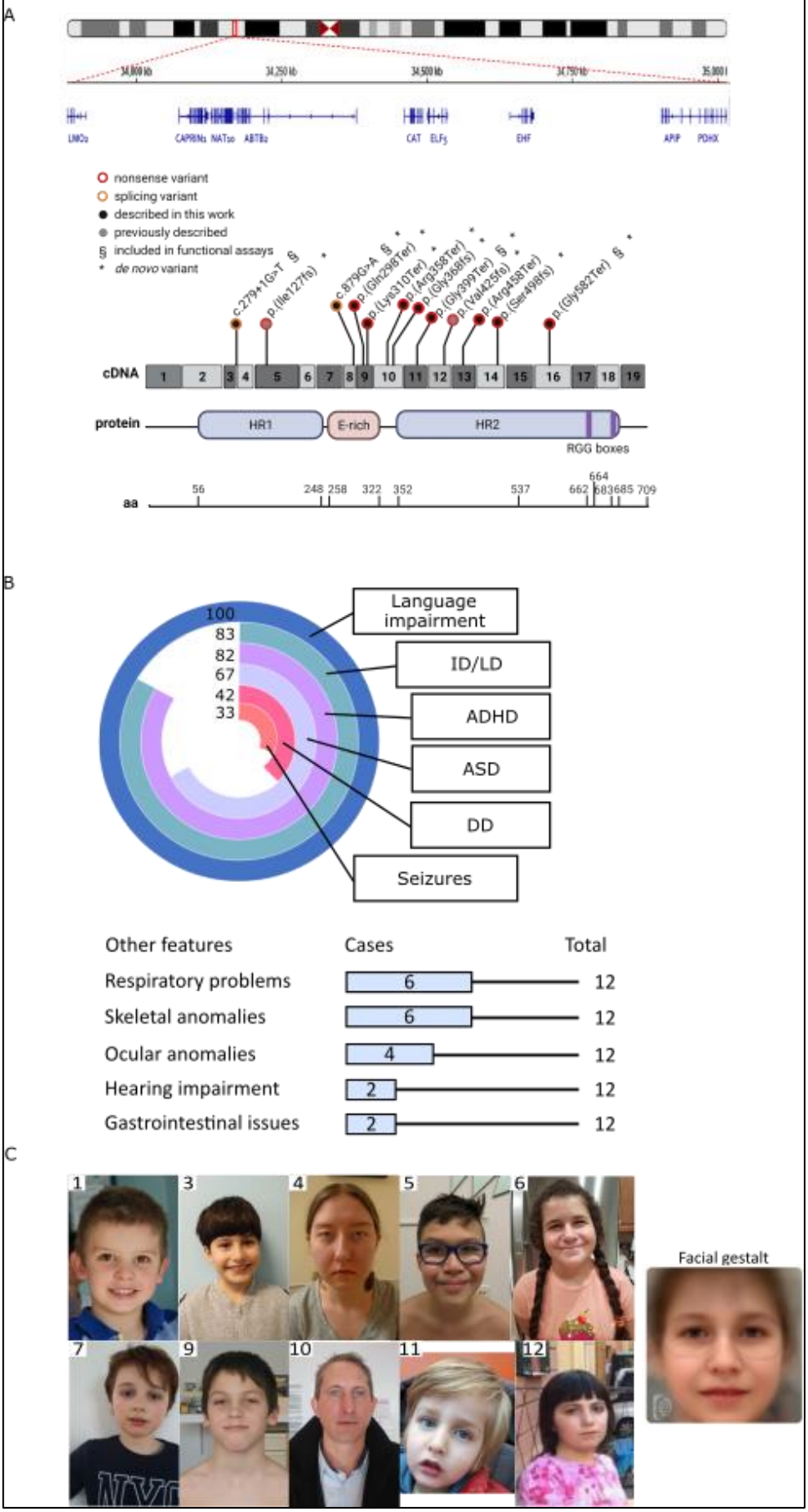

A
nonsense variant
splicing variant
described in this work
previously described
§ included in functional assays
* de novo variant
cDNA
protein
HR1
E-rich
HR2
RGG boxes
aa
B
Language impairment
ID/LD
ADHD
ASD
DD
Seizures
100
83
82
67
42
33
Other features
Cases
Total
Respiratory problems 6 12
Skeletal anomalies 6 12
Ocular anomalies 4 12
Hearing impairment 2 12
Gastrointestinal issues 2 12
C
Facial gestalt

Figure 2

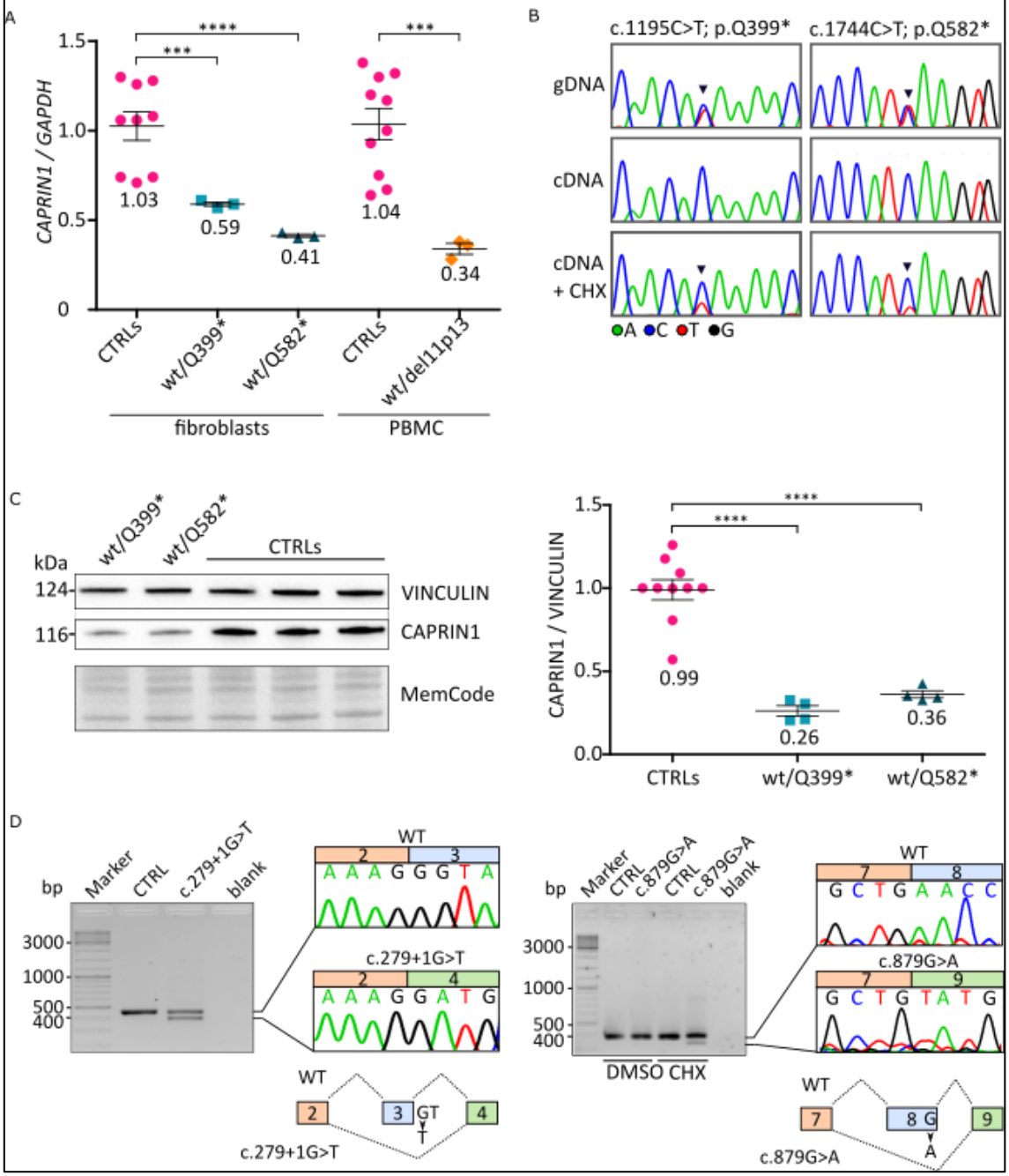

A
CAPRIN1 / GAPDH
1.5
1.0
0.5
0
****
***
***
1.03
0.59
0.41
1.04
0.34
CTRLs
wt/Q399*
wt/Q582*
CTRLs
wt/del11p13
fibroblasts
PBMC
B
c.1195C>T; p.Q399*
c.1744C>T; p.Q582*
gDNA
cDNA
cDNA + CHX
A C T G
C
kDa
wt/Q399*
wt/Q582*
CTRLs
124
116
VINCULIN
CAPRIN1
MemCode
CAPRIN1 / VINCULIN
1.5
1.0
0.5
0.0
****
****
0.99
0.26
0.36
CTRLs
wt/Q399*
wt/Q582*
D
bp
Marker
CTRL
c.279+1G>T
blank
3000
1000
500
400
WT
2
3
A A A G G G T A
c.279+1G>T
2
4
A A A G G A T G
WT
2
3 GT
T
4
c.279+1G>T
bp
Marker
CTRL
c.879G>A
CTRL
c.879G>A
blank
3000
1000
500
400
DMSO CHX
WT
7
8
G C T G A A C C
c.879G>A
7
9
G C T G T A T G
WT
7
8 G
A
9
c.879G>A

Figure 3

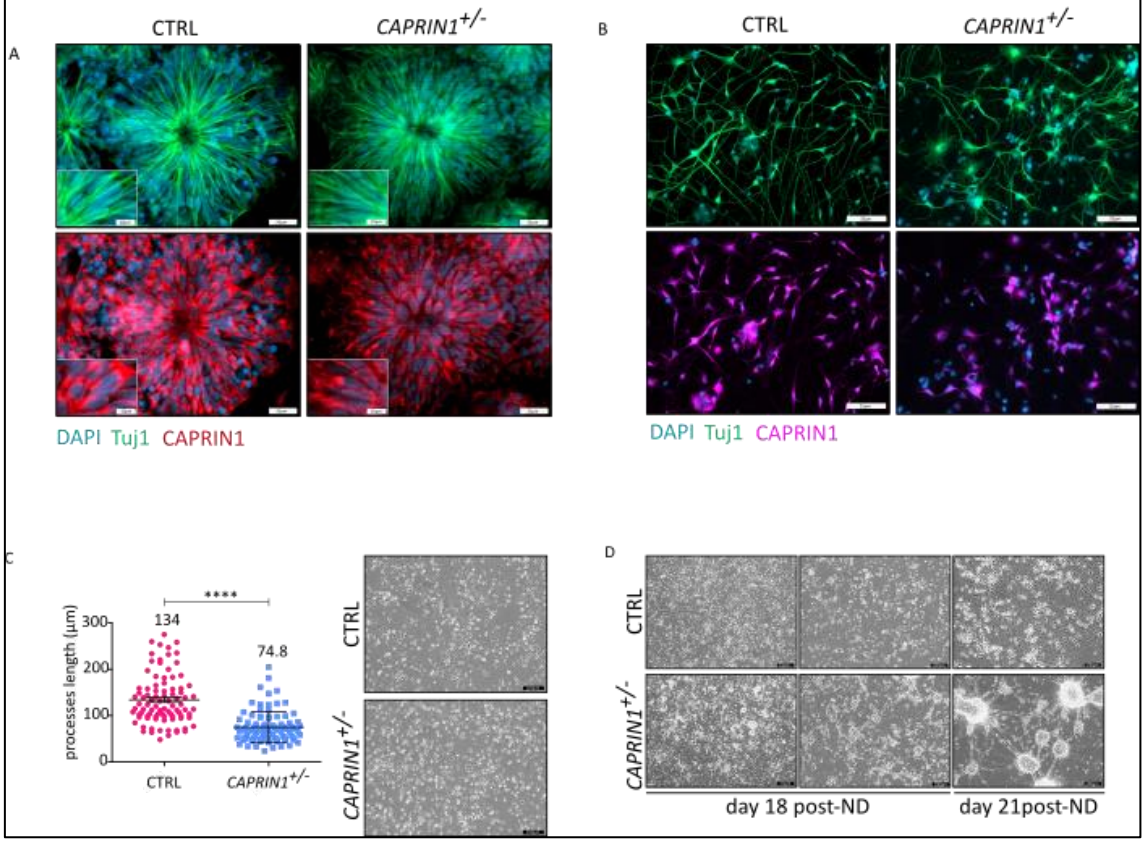


Figure 4

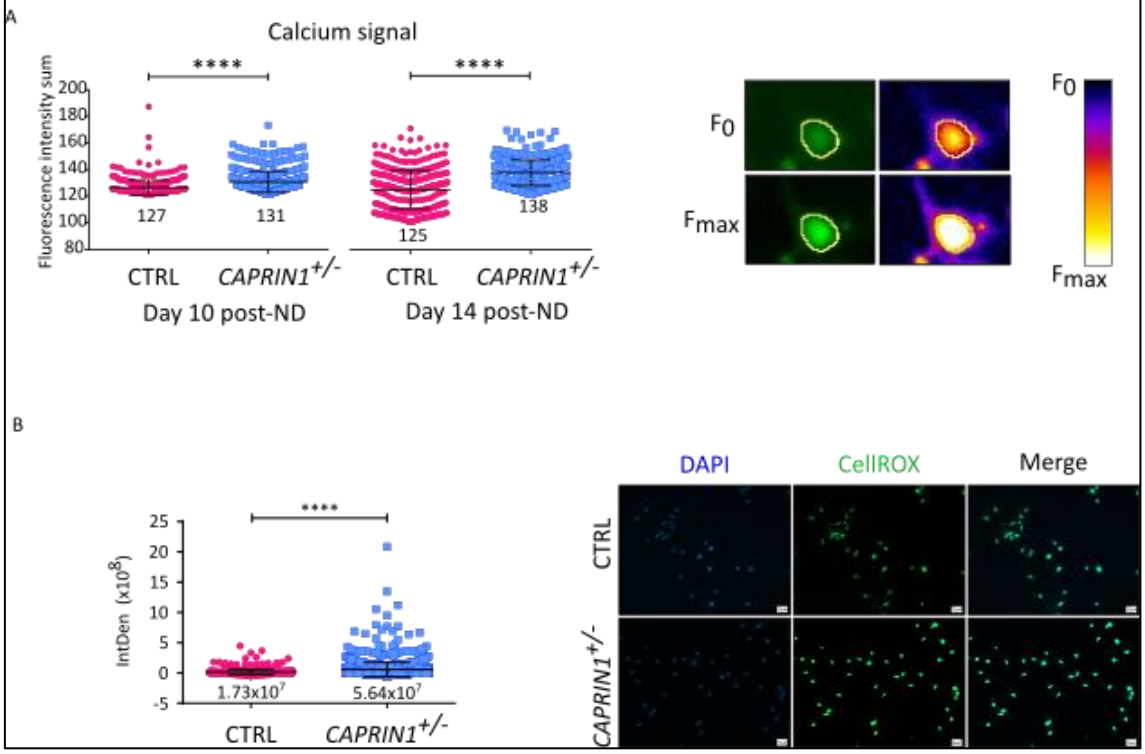


Figure 5

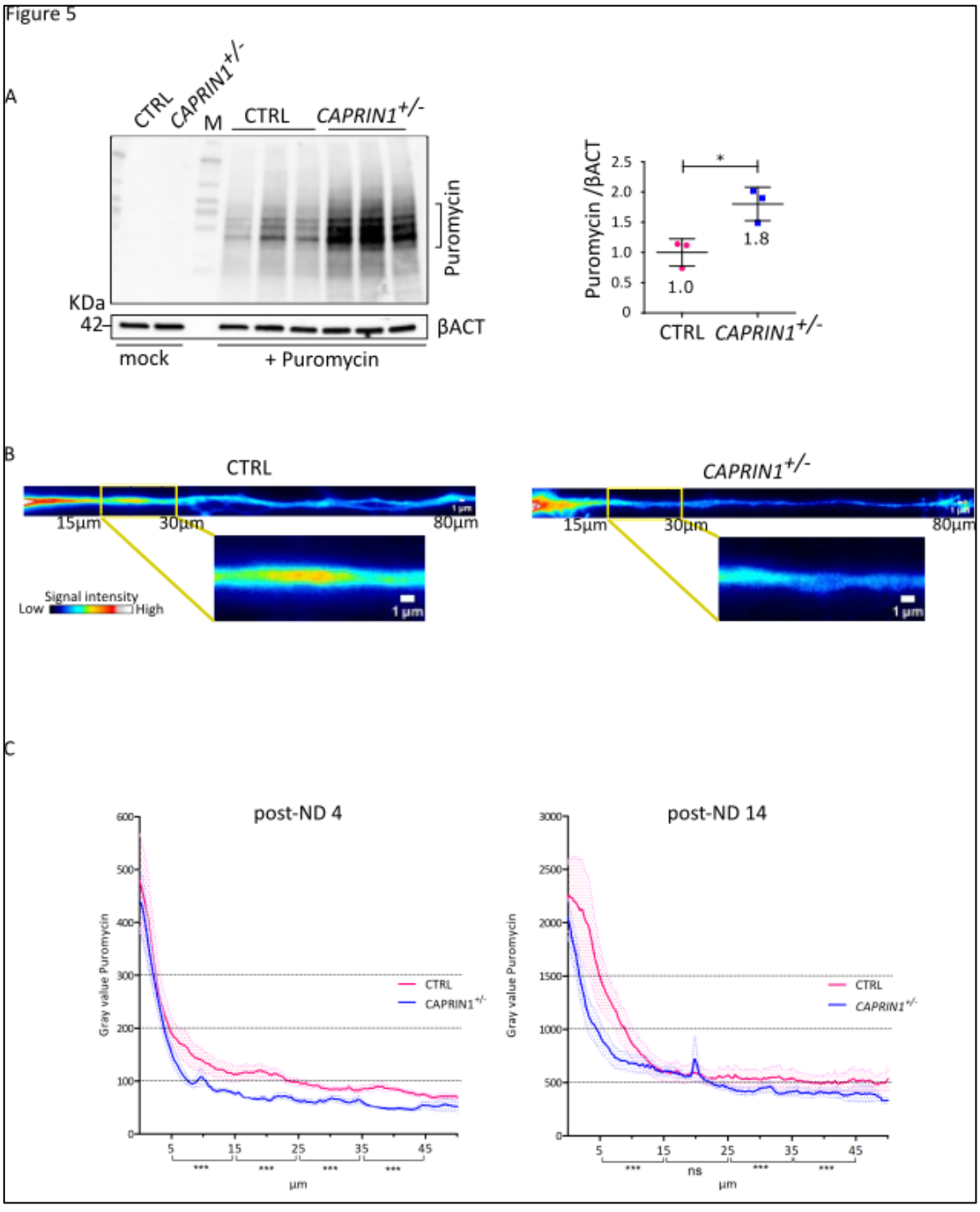

A
CTRL
CAPRIN1+/-
M
CTRL
CAPRIN1+/-
Puromycin
KDa
42
βACT
mock
+ Puromycin
Puromycin /βACT
2.5
2.0
1.5
1.0
0.5
0
*
1.8
1.0
CTRL
CAPRIN1+/-
B
CTRL
CAPRIN1+/-
15µm
30µm
80µm
Signal intensity
Low
High
1 µm
C
post-ND 4
post-ND 14
Gray value Puromycin
CTRL
CAPRIN1+/-
***
ns
µm

Figure 6

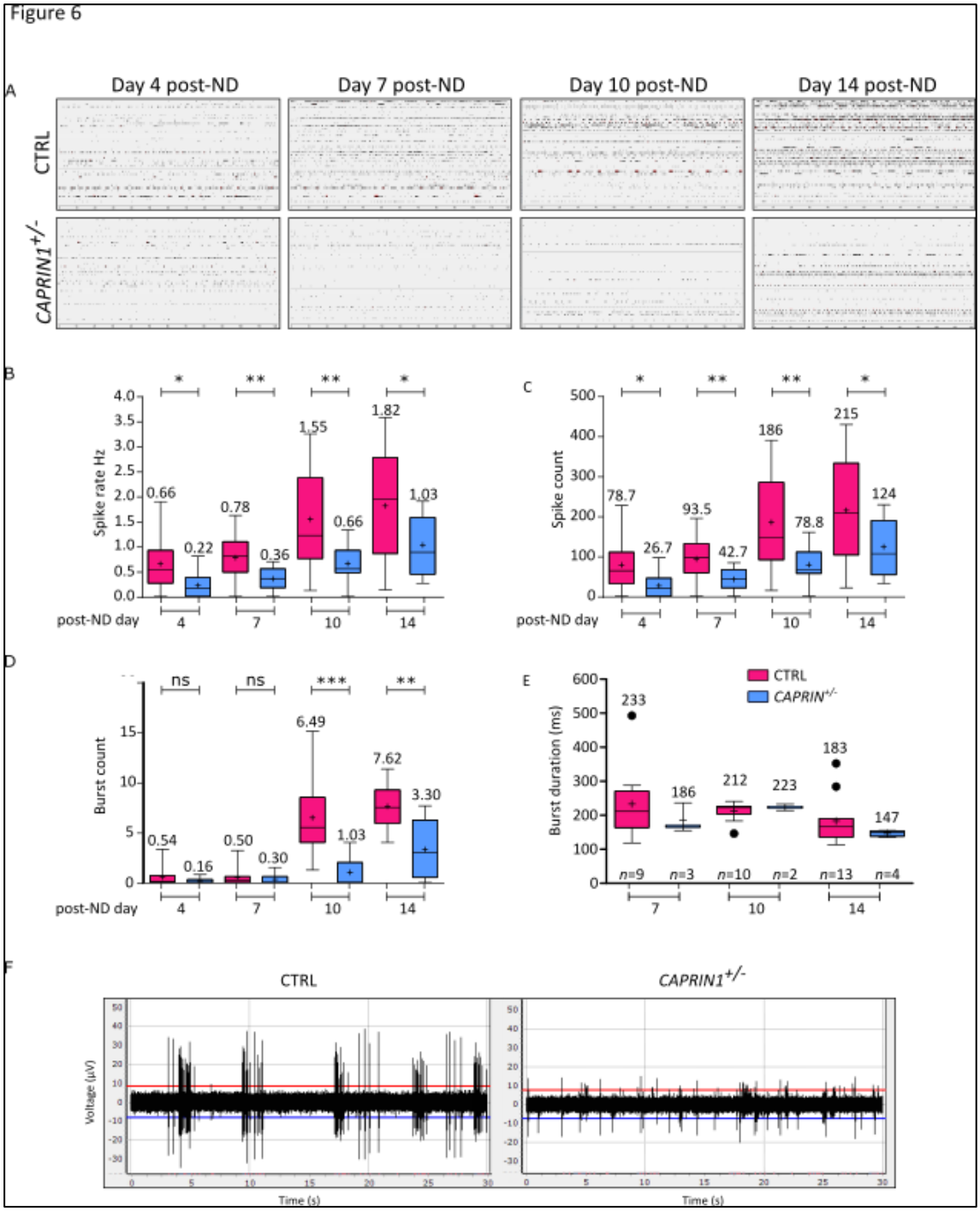

A
Day 4 post-ND
Day 7 post-ND
Day 10 post-ND
Day 14 post-ND
CTRL
CAPRIN1+/-
B
Spike rate Hz
post-ND day
0.66
0.22
0.78
0.36
1.55
0.66
1.82
1.03
C
Spike count
78.7
26.7
93.5
42.7
186
78.8
215
124
D
Burst count
ns
ns
***
**
0.54
0.16
0.50
0.30
6.49
1.03
7.62
3.30
E
Burst duration (ms)
CTRL
CAPRIN+/-
233
186
212
223
183
147
n=9
n=3
n=10
n=2
n=13
n=4
F
CTRL
CAPRIN1+/-
Voltage (µV)
Time (s)

Figure 7

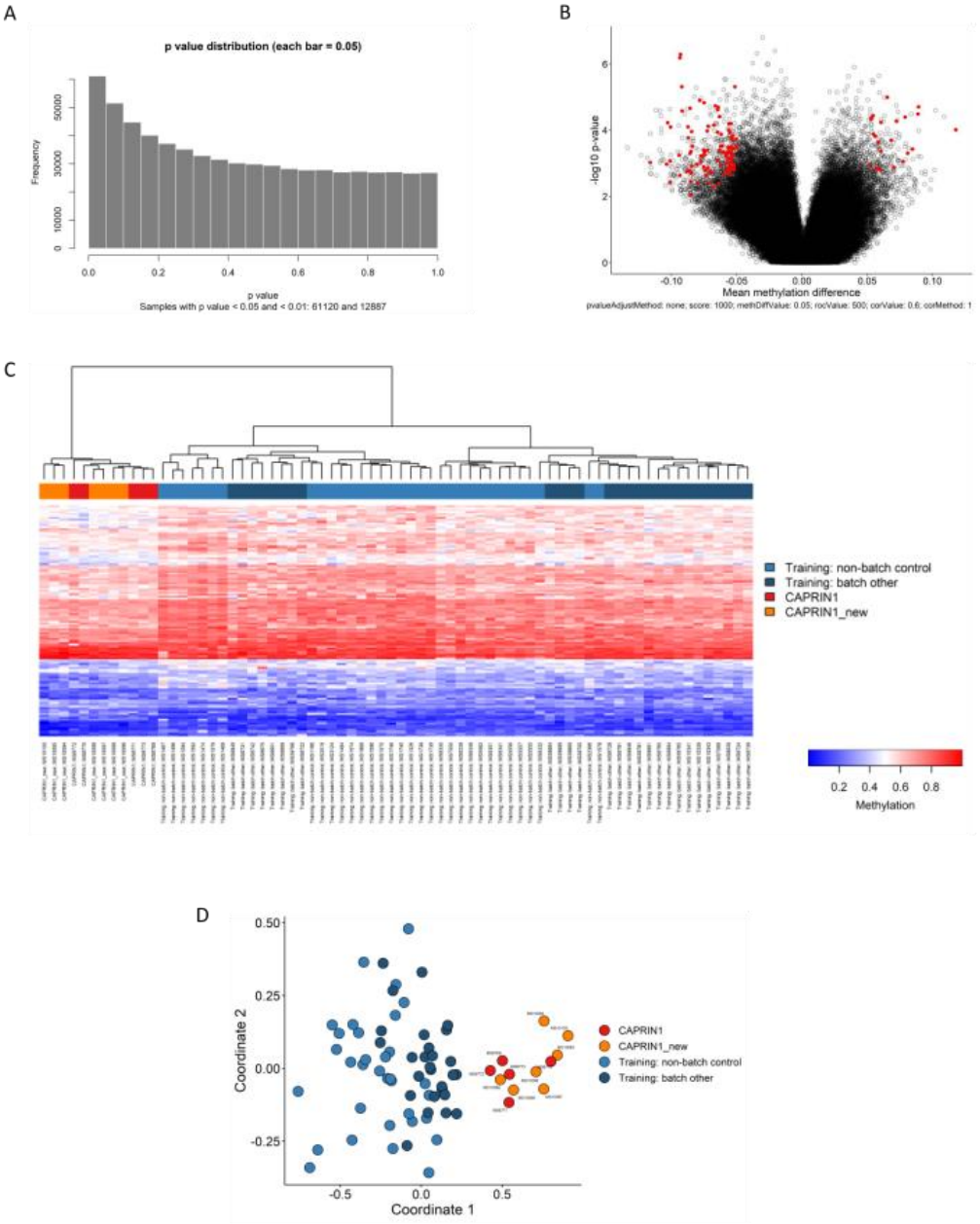
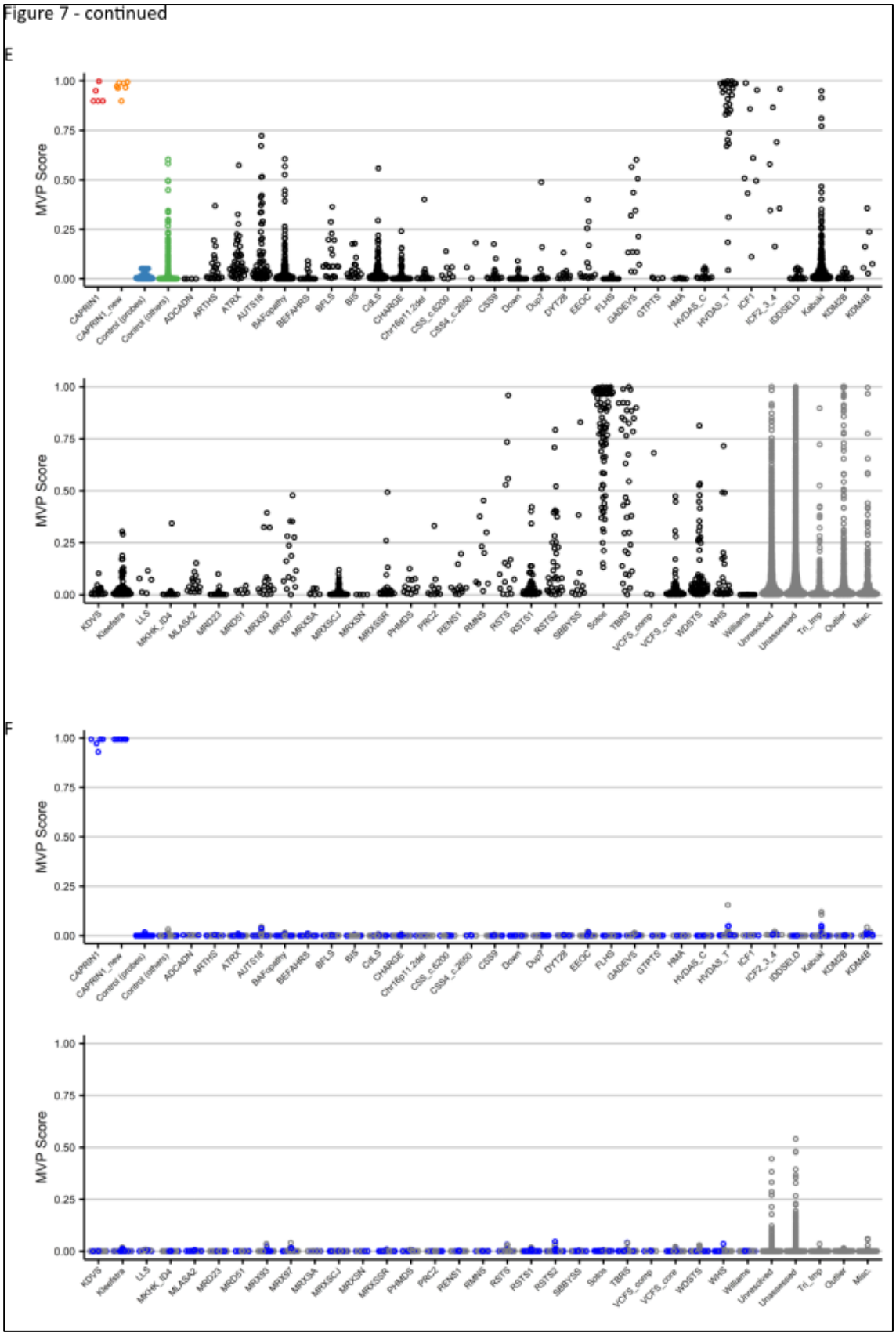

A
p value distribution (each bar = 0.05)
Frequency
p value
Samples with p value < 0.05 and < 0.01: 61120 and 12887
B
-log10 p-value
Mean methylation difference
C
Training: non-batch control
Training: batch other
CAPRIN1
CAPRIN1_new
Methylation
D
Coordinate 1
Coordinate 2
CAPRIN1
CAPRIN1_new
Training: non-batch control
Training: batch other

Figure 7 - continued



2.2 RPH3A is associated with a clinically variable neurodevelopmental disorder

N-methyl-D-aspartate ionotropic glutamate receptors (NMDARs) drive key cellular events involved in neuronal differentiation and synapse formation.[113] These processes are correlated to dynamic modifications of NMDAR subunit composition, including the developmental switch from GluN2B- to GluN2A-containing receptors.[114] Several pathogenic variants involving NMDAR subunits (e.g., GRIN proteins) have been identified and associated with a wide spectrum of neurodevelopmental disorders (NDDs), including epilepsy and autism spectrum disorder (ASD).[5,115–120] Among these, mutations affecting the *GRIN2A* gene (MIM:138253) coding for the GluN2A regulatory subunit represent the most common event,[121,122] leading to NMDAR gain of function or loss of function and causing a wide spectrum of epilepsies in the context of different types of neurodevelopmental disorders.[123–125]

Synaptic transmission and plasticity at glutamatergic synapses depend on the correct synaptic localization of GluN2A-containing NMDARs.[126] Accordingly, the involvement of the GluN2A subunit in neurodevelopmental disorders and epilepsy (MIM: 245570) has been correlated not only to aberrant receptor activation, but also to altered synaptic trafficking and stabilization and balance with other NMDAR subunits for an adequate receptor functioning[127,128]. As a direct consequence, protein-protein interactions at the GluN2A C-terminal domain (CTD) driving GluN2A synaptic retention have been widely investigated.[127]

GluN2A binds to scaffolding proteins of the PSD-MAGUK family (e.g., PSD-95) and this interaction anchors the receptor at the postsynaptic membranes.[127] Notably, mutations in *DLG4* (MIM: 602887), the gene coding for PSD-95, have been associated with intellectual disability (ID), ASD, and epilepsy (MIM:618793).[129,130] We identified Rabphilin 3A, encoded by *RPH3A* gene (MIM: 612159), as a novel GluN2A synaptic partner needed to stabilize GluN2A/PSD-95 complex at the postsynaptic density[131]. Disruption of the Rph3A/GluN2A/PSD-95 complex reduces GluN2A synaptic retention and promotes NMDARs endocytosis.[127,131,132] *GRIN2A* and *DLG4* have been associated with rare autosomal dominant NDDs and epilepsy[124,129,130], but so far *RPH3A* has not been genetically associated with NDDs. Here, we provide genetic and functional evidence for a role of *RPH3A* in NDD risk.

Materials and methods

Primary neuronal culture

Hippocampal neurons were obtained from 18-day-old embryos of rats. The embryos were rapidly decapitated before the removal of tissues. Hippocampus was rapidly dissected under sterile conditions, kept in cold Hanks' balanced saline solution (HBSS; 4°C) with high glucose, and

then digested with papain (0.5mgml−1) dissolved in HBSS plus DNAse (0.1 mg/ml). Isolated cells were plated at a final density of 300 cells/mm^2 on plastic dishes.

Electrophysiology

Patch electrodes were fabricated from thick borosilicate glasses (Hilgenberg, Mansifield, Germany), pulled to a final resistance of 5-6 MΩ . Patch clamp recordings were performed in whole cell configuration using a Multiclamp700-B amplifier connected to a Digidata 1440 and governed by the pClamp10 software (Axon Instruments, Molecular Devices, Sunnyvale, CA, USA). Experiments were performed at room temperature (22–24°C) in whole cell configuration and acquired with sample frequency of 10 kHz. Recordings with leak current > 100pA or series resistance > 20 MΩ were discarded. Analysis was performed with Clampfit software (Axon Instruments). Whole cell voltage-clamp recordings were performed at a holding potential of -70 mV. The superfused extracellular bath solution contained: 2 mM $CaCl_2$, 130 NaCl, 2 mM $MgCl_2$, 10 mM Hepes, 10 mM glucose and 4 mM KCl (pH 7.4). The internal solution contained: 90 mM CsCl, 8 mM NaCl, 20 mM TEACl, 10 mM EGTA, 10 mM glucose, 1 mM $MgCl_2$, 4 mM ATP, 0.5 mM GTP and 15 mM phosphocreatine (pH 7.4 adjusted with CsOH). For blocking synaptic currents due to the activation of glutamatergic AMPARs and GABAergic synapses ($GABA_A$ receptors), we added respectively 6,7-dinitroquinoxaline-2,3-dione, DNQX (20μM, Sigma-Aldrich, St Louis, MO, USA) and picrotoxin (100μM, Sigma-Aldrich). Tetrodotoxin (TTX, 0.3μM, Tocris Bioscience, Bristol, UK) was added to block voltage-gated Na^+ channels. The neuron was constantly superfused through a gravity system that allowed a rapid change (50–60ms) of the solutions and the recording of I_{NMDA} following NMDA 50 (μM) and Ifenprodil (150 nM) administration. During the experiments, NMDA was administered to neurons until I_{NMDA} reached a steady-state value. I_{NMDA} charge was measured as the area under 8s of whole-cell current. GluN2B-mediated component was calculated as the difference between total NMDA current and GluN2A-dependent current measured in the presence of Ifenprodil (150 nM).

Immunocytochemistry (ICC)

For colocalization and morphological studies, transfected hippocampal neurons were fixed for 10 min at room temperature (RT) in 4% paraformaldehyde (PFA) and 4% sucrose in Dulbecco's phosphate buffered saline (PBS). Coverslips were then washed with PBS, permeabilized with 0.1% Triton X-100 in PBS for 15 min at RT and blocked for 30 min at RT with 5% bovine serum albumin (BSA) in PBS. Cells were then incubated with primary antibodies in 5% BSA-PBS overnight at 4 °C in a humidified chamber. After washes with PBS, the cells were incubated with the fluorophore-conjugated secondary antibodies in 5% BSA-PBS for 1 h at RT in a humidified chamber protected from light. The incubation was followed by washes with PBS and mounting

onto glass slides using Fluoroshield mounting medium (Sigma-Aldrich). For the surface staining assays, cells were not permeabilized and were labeled with primary antibody for extracellular epitopes overnight at 4 °C. The coverslips were then washed in PBS and a secondary antibody conjugated to Alexa Fluor dye was used.

Confocal imaging

Images were taken using LSM900 (Zeiss) confocal microscopes with 63X objective at 0,09 um pixel size and analysed using ImageJ software. Cells were chosen randomly for quantification from different coverslips from independent experiments and images were acquired using the same settings/laser power. For spine morphology, Z-stack images were taken with a Z-step of 0.35 µm. Analysis of dendritic spine morphology was performed with ImageJ. For each dendritic spine, length, head and neck width were measured, which was used to classify dendritic spines into three categories (thin, stubby and mushroom). In particular, the length and the ratio between the width of head and the width of neck (Wh/Wn) were used as parameters for the classification as follows: protrusions having a length of more than 3 µm were considered as filopodia, the others as spines; spines with a Wh/Wn ratio bigger than 1.7 were considered mushrooms; spines with a Wh/Wn ratio smaller than 1.7 were divided in stubby, if shorter than 1 µm, and thin if longer than 1 µm. Protrusions with length over 5 µm were excluded from the analysis. For each neuron a total length of about 150-200 µm of proximal dendrites and a number of about 100-150 dendritic spines were considered. For super-resolution assays, images were taken using the Airyscan mode on a LSM900 confocal microscope (Zeiss) with a 63X objective at a 0.04 µm pixel size.

Analysis of GCaMP7 signals

Calcium imaging experiments using pCAG_Xph20-GCaMP7f calcium indicator were performed in primary hippocampal neuronal cultures at DIV16. pCAG_Xph20-GCaMP7f was a gift from Matthieu Sainlos (Addgene plasmid #135539; http://n2t.net/addgene:135539; RRID: Addgene_135539). Live-cell imaging acquisition was performed using LSM900 (Zeiss) confocal microscopes with 63X objective. Dendrites were imaged in artificial cerebral spinal fluid (ACSF) at 37ºC in an atmosphere of 5% CO_2 at maximum scanning speed with 6.7 zoom (scan time around 35 ms, resolution: 0.07 nm/pixel) and time-lapse images were taken for 3 minutes at 10 Hz. For quantitative analysis, dendritic spines were chosen as ROI, and the GCaMP7f fluorescence was plotted overtime using ImageJ.

Results

Identification of patients with variants in *RPH3A* gene

We examined trio-based exome sequencing (ES) data from 538 individuals diagnosed with NDDs in local hospitals, within the Autism Sequencing Consortium project[5,26] for any rare *RPH3A* variants not seen in gnomAD[133] or *de novo.* Written informed consent for genetic testing and permission for publication of the clinical data was obtained from all individuals and/or their legal representatives by the referring physicians according to the guidelines of the ethics committees and institutional review boards of the respective institutes. We identified a male individual (case #1) with a *de novo RPH3A* variant (NM_001143854.2: c.1853A>G, p.Asn618Ser) and high functioning ASD, mild ID, dysgraphia, hyperactivity, increased body weight and abnormal facial features (Supplementary note: clinical reports). Querying GeneMatcher[40] ,we identified a male individual also carrying a *de novo RPH3A* missense variant c.1349C>G (p.Thr450Ser) (case #2; ClinVar SCV000265605.3) and showing a different neurodevelopmental phenotype, with drug-resistant epilepsy and moderate ID. Array-CGH and Fragile-X syndrome testing were negative for all individuals.

Finally, by searching the Exomiser-based candidate variants from the individuals with clinical indications 'intellectual disability' (n = 6203) and 'epilepsy plus other features' (n = 1362) in the 100,000 Genomes Project[41], we identified four additional males with missense variants in *RPH3A*, namely c.626G>A (p.Arg209Lys) (case #3), c.705G>T (p.Arg235Ser) (case #4) and c.1524G>C (p.Gln508His) (case #5) (Fig. 1A). Parents of cases #3 and #4 were not available for segregation analysis. The unaffected mother of case #5 was found not to carry the (p.Arg235Ser) variant. Clinically, cases #2, #3 and #5 showed drug-resistant epilepsy and moderate ID, similar to case #2. Case #4 showed childhood autism spectrum disorder, intellectual disability, and obesity, similar to case #1.

The Exomiser-based analysis revealed an additional male with the c.444G>T change, predicted to cause a Glu148Asp substitution. Indeed, the variant hit the last nucleotide of exon 7, caused an alteration of splicing with 72 bp intron retention and a premature stop codon (Fig. S1). Because *RPH3A* is unlikely to be an haploinsufficient gene (*see discussion*, GnomAD pLi=0 [133,134]), we excluded this case from our survey.

All the identified missense variants were absent from the Genome Aggregation Database (GnomAD, v2.1.1[133]), hit evolutionary conserved amino acid residues that had a mutation tolerance score from neutral (Lys209 and Arg235) to intolerant (Thr450 and Gln508) and highly intolerant (Arg618) using MetaDome[109] (Fig. 1B,C); bioinformatic predictions were not concordant in the pathogenicity score (table S1): for instance CADD score were above 20 for p.Thr450Ser, p.Gln508His and p.Asn618Ser. Of note, c.1853A>G (p.Asn618Ser) occurred five

bases upstream of the intron 20 splice donor site. We excluded a possible impact on mRNA processing by pSPL3 minigene assay (Fig. S2).[135]

In silico evaluation of identified variants

To explore the functional effect of the *RPH3A* variants, we modelled *in silico* and *in vitro* two variants localized in key domains of the protein and associated with different phenotypes: p.Thr450Ser (hereby indicated as *h*Rph3A^{T450S}) (identified in case #2) and p.Asn618Ser (hereby indicated as *h*Rph3A^{N618S}) (identified in case #1). Both changes affect residues located within the Rph3A C2A and C2B domains, which play a fundamental role in the regulation of activity-dependent protein function at the postsynaptic compartment[127]. Using two crystallographic structures of *Rattus norvegicus* Rph3A (PDB ID: 4LT7 for the C2A domain; PDB ID: 2CM6 for the C2B domain[136]);[137], we first built three-dimensional models of *h*Rph3A^{WT}, and [3132]then generated the *h*Rph3A^{T450S} and *h*Rph3A^{N618S} mutants. Moreover, the experimentally solved structure of *Mus* musculus Rph3A (PDB ID: 2K3H)[138] was carefully considered to evaluate the side-chain orientations of residues near to Thr450 and Ca^{2+} ions. Multiple alignment information is reported in figure S3. The 2K3H structure contains 20 conformers solved by NMR[138]. We observed that the carboxyl group of Glu472 is closed to Asn481 in only 4 conformers, whereas the carboxyl group of Glu472 is closed to Thr437 in the remaining 16, similarly to what occurs in the 2LT7 crystallographic structure. *h*Rph3A^{WT} showed the same local interaction network of the mouse and rat proteins: the carboxyl group of Glu485 forms a hydrogen bond with the hydroxy group of Thr450, and the backbone of Glu485 cooperates with Leu427, Asp429, Asp486 and Asp487 in binding one Ca^{2+} ion.

The impact of the Thr-to-Ser substitution on the overall *h*Rph3A^{T450S} stability is predicted to be minimal, as the total free energy increased of approx. 1.3 kcal/mol. On the other hand, a rotamer comparative analysis between the *h*Rph3A^{WT} and *h*Rph3A^{T450S} structures suggested a different local stability of Thr and Ser rotamers positioned at 450. The energy variation associated to different rotameric configurations is related to a significant modification of the local intramolecular interaction network. Thr is characterized just by one stable rotamer, forming a hydrogen bond with the side chain of Glu485; differently, the most stable Ser rotamer, with a minimum energy value lower (approx. 3.5 kcal/mol) than other Ser rotamers, is not able to interact with Glu485. This conformation of the Ser side chain strongly impacts the interaction network of the residues directly interacting with Ca^{2+} ions (Fig. 2A). The change of this interaction network was expected to alter the C2A affinity to inositol-1,4,5-trisphosphate (IP3), inhibiting protein recruitment in membrane. In fact, C2A of Rph3A^{WT} binds IP3 in a Ca^{2+}-

dependent manner, with an affinity in the micromolar range in presence of saturating concentration of Ca^{2+}.[138] In contrast, the C2B domain of Rph3A^{WT} binds IP3 in a Ca^{2+}-independent manner, with an affinity in the millimolar range.[138] Similar to *h*Rph3A^{T450S}, *h*Rph3A^{N618S} does not significantly reduce protein stability, increasing the total free energy of just ~0.6 kcal/mol. However, this mutation locally decreases the hydrophobic surface area of ~10 $Å^2$ (Fig. 2A), and the overall protein volume of ~26 $Å^3$. These effects depend on the interaction network of Glu616 that establishes a hydrogen bond with Arg573 and residue 618 in both *h*Rph3A^{WT} and *h*Rph3A^{N618S}. While Ser618 maintains the same interaction network of Asn618, its side chain is shorter, causing a contraction of the total protein volume.

Rph3A variants cause increase of total and GluN2A-dependent NMDA currents

To evaluate whether the p.Thr437Ser and p.Asn605Ser substitutions could affect NMDAR function, we examined NMDAR currents (I_{NMDA})[139,140] via whole-cell patch clamp recordings on rat hippocampal neuronal cultures[34,35]. Specifically, we compared the I_{NMDA} amplitudes by bath application of exogenous NMDA in RFP-Rph3A^{WT} versus RFP-Rph3A^{T437S} or RFP-Rph3A^{N605S} transfected neurons (Fig. 3A). Neurons were transfected at day in vitro 8 (DIV8) and all recordings were performed at DIV16. Whole-cell recordings were performed by holding membrane potential at -70 mV. Exogenous application of NMDA (50 µM) induced a mean I_{NMDA} amplitude of 283.3 ± 28.23 pA in neurons expressing RFP-Rph3A^{WT}, that significantly increased in neurons expressing either mutant (RFP-Rph3A^{T437S}, 602.75 ± 68.79 pA, $p<0.0001$; RFP-Rph3A^{N605S}, 590.33 ± 79.03 pA, $p<0.01$) (Fig. 3C).

In mature hippocampal neurons NMDA whole-cell currents depend mostly on GluN2A and GluN2B components[27]. Therefore, we investigated the specific contribution of GluN2A to the total currents by using ifenprodil (Fig. 3B), a non-competitive GluN2B- inhibitor[26]. In the presence of ifenprodil (150 nM), we observed a 45% inhibition of I_{NMDA} in neurons overexpressing RFP-Rph3A^{WT}, in line with previous observations[27]. In contrast, a lower reduction of I_{NMDA} amplitude was observed in neurons expressing the RFP-Rph3A^{T437S} or Rph3A^{N605S} mutants (24% and 21%, respectively). Particularly, the mean GluN2A-mediated currents were significantly higher in neurons expressing RFP-Rph3A^{T437S} or RFP-Rph3A^{N605S} when compared to control neurons (I_{NMDA} RFP-Rph3A^{WT}: 157.03 ± 27.44 pA; RFP-Rph3A^{T437S}: 460.01 ± 65.63 pA, $p< 0.001$; RFP-Rph3A^{N605S}: 463.89 ± 77.28 pA, $p< 0.0001$) (Fig. 3C), whereas the mean GluN2B-dependent current did not show differences between groups (Fig. 3C). To further validate these results and exclude any effects produced by changes of the deactivation kinetic of I_{NMDA}, we next assessed the charge underlying I_{NMDA} (Fig.3D). We found that I_{NMDA} charge in neurons expressing

RFP-Rph3A^{T437S} or RFP-Rph3A^{N605S} was increased compared to neurons overexpressing the WT protein (RFP-Rph3A^{WT}: 1.95 ± 0.23 nC; RFP-Rph3A^{T437S}: 4.40 ± 0.48 nC, p< 0.01; RFP-Rph3A^{N605S}: 3.65 ± 0.60 nC, p<0.0001). The I_{NMDA} charge dependent by GluN2A subunits was significantly higher in neurons expressing RFP-Rph3A^{T437S} and RFP-Rph3A^{N605S} mutants compared to control neurons (RFP-Rph3A^{WT}: 1.12 ± 189.01 nC; RFP-Rph3A^{T437S}: 3.11 ± 0.40 nC, p<0.001; RFP-Rph3A^{N605S}: 2.56 ± 0.56 nC; p<0.05) (Fig. 3D). Overall, these data show an increased activity of GluN2A-containing NMDARs in Rph3A^{T437S} and Rph3A^{N605S} mutants, suggesting augmented levels of GluN2A-containing NMDARs at the membrane.

NMDAR-GluN2A localization and availability is affected by *Rph3A* variants

To directly test this hypothesis, we evaluated GluN2A surface levels in the dendrites of rat primary hippocampal neurons. We overexpressed RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} in hippocampal neurons by transfection at DIV8, and the endogenous GluN2A surface level was analysed at DIV16 normalized to the total level of GluN2A. As shown in Fig. 4A-B, overexpression of RFP-Rph3A^{T437S} induced a significant increase of GluN2A surface levels in dendrites, compared with neurons overexpressing RFP-Rph3A^{WT} (p< 0.01, Rph3A^{T437S} vs Rph3A^{WT}) while RFP-Rph3A^{N605S} did not modify GluN2A surface availability (p= 0.68, Rph3A^{N605S} vs Rph3A^{WT}).

NMDARs are highly mobile and can be localized both at synaptic and extrasynaptic sites, depending upon subunit composition. Stimulation of synaptic NMDARs promotes cell survival, whereas the stimulation of extrasynaptic receptors results in neurotoxicity and can trigger intracellular pathways that lead to cell death.[113,141] To test if the *RPH3A* variants alter NMDAR synaptic retention, we examined the colocalization of GluN2A with PSD-95, the main scaffolding protein of the excitatory postsynaptic density involved in the physiological synaptic localization of NMDARs.[113,127] RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} were overexpressed in hippocampal neurons at DIV8 and then immunocytochemistry assay for GluN2A and PSD-95 was performed at DIV16. Confocal microscopy with Airyscan super-resolution module was used to identify co-localization puncta. As shown in Fig. 4C-D, the percentage of GluN2A colocalizing with PSD-95 was decreased in neurons expressing either mutant (p< 0.0001, Rph3A^{N605S} vs Rph3A^{WT}; p< 0.01, Rph3A^{T437S} vs Rph3A^{WT}).

Rph3A variants affect postsynaptic calcium currents

To assess whether reduced GluN2A synaptic localization in Rph3A mutant neurons was associated with altered postsynaptic calcium levels, we used the novel genetically encoded

calcium-sensing probe pCAG_Xph20-GCaMP7f containing the Xph20 moiety for specific and efficient binding to endogenous PSD-95. This approach allows for the specific evaluation of calcium transients at dendritic spines using live imaging. As shown in Fig. 4E, neurons overexpressing Rph3A^{N605S} or Rph3A^{T437S} were characterized at resting by an alteration of postsynaptic calcium transients. Rph3A^{T437S} was characterized by lower area under the curve of calcium signals compared to neurons expressing Rph3A^{WT} (Fig. 4F, $p< 0.01$, Rph3A^{T437S} vs Rph3A^{WT}) while Rph3A^{N605S} showed a reduction of number of calcium events (Fig. 4G; $p< 0.01$, Rph3A^{N605S} vs Rph3A^{WT}).

Rph3a^{T437S} affects dendritic spines morphology

Rph3A can interact with MyosinVA (MyoVA).[142] Since MyoVA acts as motor protein transporting GluA1-containing AMPARs in dendritic spines,[143] Rph3A mutants might alter the membrane insertion of AMPARs during synaptic plasticity events. Therefore, we quantified the level of AMPARs on the membrane by analyzing the GluA1 subunit surface availability in dendrites of hippocampal neurons. As shown in Fig. 5A-B, analysis of GluA1 surface levels did not show significant differences between the three experimental groups, suggesting that *RPH3A* mutations impact only NMDA-type ionotropic glutamate receptors. Excitatory neurotransmission strictly modulates dendritic spine morphology and density, reflecting specific states of the synapse. Notably, Rph3A synaptic activity has been closely correlated to modifications of dendritic spine density.[127,131,132] Based on these data, we investigated the impact of Rph3A on dendritic spine formation and maturation by examining their density and morphology. To this end, hippocampal primary neurons were co-transfected with GFP and RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} and RFP-Rph3A^{T437S}, and then analyzed for evaluation of dendritic spine density and shape. We did not observe any significant change in dendritic spine density (Fig. 5C-D). For a more detailed morphological analysis, the dendritic spine length, head and neck width were measured, and spines were categorized within the filopodia, mushroom, stubby and thin classes (Fig. 5E-H). No differences were found in length of spines (Fig. 5E) and percentage of filopodia (Fig. 5G). However, a significant reduction of head width was associated with the Rph3A^{T437S} mutant ($p< 0.05$, Rph3A^{T437S} vs Rph3A^{WT}) (Fig. 5F). Accordingly, the number of thin spines of neurons overexpressing Rph3A^{T437S} were significantly increased compared with the number of thin spines of neurons overexpressing Rph3A^{WT} ($p< 0.05$, Rph3A^{T437S} vs Rph3A^{WT}) (Fig. 5H). Overall, these results indicate that only the p.Thr437Ser change induces spine morphology changes, at least in the present experimental conditions.

Discussion

Our results indicate that the p.Thr437Ser amino acid substitution leads to widespread alterations at hippocampal neurons including aberrant activation of NMDARs, alterations of GluN2A synaptic localization and surface levels and dynamics of calcium transients at dendritic spines. These molecular/functional events are also associated to the presence of thinner and longer dendritic spines. In addition, structural analysis showing an impact of p.Thr437Ser on the interaction network of the residues directly interacting with Ca^{2+} ions fully support imaging data. Similar but less strong modifications were observed with the $Rph3A^{N605S}$ mutant not displaying modifications of GluN2A surface levels and morphological changes of spines. Importantly, both $Rph3A^{T437S}$ and $Rph3A^{N605S}$ did not induce alterations of AMPARs, suggesting a specific effect on NMDARs. Our data agree with previous studies indicating that a balanced Rph3A activity at the synaptic NMDAR complex is needed for the induction of synaptic plasticity and NMDAR-dependent behaviors[127,132]. The altered synaptic retention of GluN2A-containing NMDARs observed here in neurons transfected with both $Rph3A^{T437S}$ and $Rph3A^{N605S}$ suggests that their stimulation can promote pathological glutamatergic events known to be mostly associated with extrasynaptic NMDARs[113,126]. Importantly, aberrant Rph3A interactions with GluN2A-containing NMDARs were previously observed in models of pathological synaptic plasticity, such as parkinsonian rats displaying a dyskinetic profile[144].

RPH3A has been only sporadically suggested to be associated with human disease. Compound heterozygous variants impacting residues 269 and 464 were reported in an individual with congenital myasthenic syndrome,[145] though the functional and clinical relevance of those changes have not been established[145] . More recently, a 2Mb-large region encompassing *RPH3A* gene has been reported as a novel epilepsy-risk region, hosting multiple epilepsy-associated SNPs affecting *RPH3A* expression.[146] In our cohort we identified two groups of cases, showing drug-resistant epilepsy and ID (cases #2, #3, #4 and #6) and high functioning ASD and learning disability (cases #1 and #5). Case #1 showed additional features compared to other reported individuals, including language and motor impairment, breathing issues and mixed hypo/hypertonia. Notably, he also had variants in the *CUL4B*, *PRKAG2* and *SCN4A* genes that could potentially explain some of his clinical features. None of these genes has been associated with seizures, confirming *RPH3A* as the best candidate for this feature.

A loss-of-function pathogenic mechanism for *RPH3A* is unlikely given haploinsufficiency scores are compatible with a recessive gene (pLI: 0; LOEUF: 0.47);[133,134] in addition, mice lacking *Rph3a* are viable and do not show gross alterations[147]. For this reason, we excluded the case carrying the c.444G>T splicing variant, as it caused the insertion of a premature stop codon. Accordingly,

our experiments suggest a functional gain of function for the two studied variants (p.Thr437Ser and p.Asn605Ser), with increased GluN2A mediated-NMDAR currents, increased extrasynaptic GluN2A availability on membrane and enhanced postsynaptic calcium transients. It is plausible to hypothesize that different missense variants may have variable impact on Rph3A function, depending on their localization along the protein. The identification of additional subjects with pathogenic variants in this gene is expected to provide a more accurate pictures of the clinical variability and genotype-phenotype correlations associated with altered Rph3a function.

In conclusion, here, we provide evidence of a causal association between missense variants in the *RPH3A* gene and an ultra-rare neurodevelopmental disorder with variable expressivity, associated with epilepsy and ID or ASD and learning disability.

Figures and tables

Figure 1. *RPH3A* variants in five individuals.

(A) Pedigrees of the five identified RPH3A variants. For T450S and N618S, parents were available, and the variants were verified as *de novo*. **(B)** Relevant selection of the multiple sequence alignment (MSA) of mammalian RPH3A orthologs (*Homo sapiens, Macaca mulatta, Mus musculus, Rattus norvegicus, Sus scrofa, Laxodonta Africana, Dasypus novemcinctus*). The MSA highlights the conservation of the R209, R235, T450, Q508 and N618 aminoacids. **(C)** Mutation tolerance at each position in RPH3A human protein as calculated by MetaDome. Below the schematic representation of RPH3A protein, based on InterPro data (http://www.ebi.ac.uk/interpro/). Rab-binding domain (aminoacids 44-160), and the two C2 domains (C2A aminoacids 392-514; C2B aminoacids 550-683) are reported. *RPH3A* (NM_001143854.2) variants described in the present work are represented along the protein with the phenotype associated (ID=intellectual disability; ASD= autism spectrum disorder).

Figure 2. *h*Rph3A^{T450S} alters Rph3A conformation *in silico*.

(A) On the top is the C2A domain of *h*Rph3A^{WT} vs. *h*Rph3A^{T450S}. Interaction network of Ca^{2+} involving Asp484, Glu485, Asp429, Asp486, and Thr450 is shown for *h*Rph3A^{WT}. Notably, network of Ca^{2+} for *h*Rph3A^{T450S} involves the same residues but not Ser450, due to the loss of hydrogen bond with the side chain of Glu485. This event has an important effect on the interaction network of the residues directly interacting with Ca^{2+} ions.

On the bottom, the C2B domain of *h*Rph3A^{WT} vs. *h*Rph3A^{N618S}. The interaction network involves Glu616, which establishes a hydrogen bond with Arg573 and residue Asn618. The network is not impaired for Ser618. However, as the side chain of Ser618 is shorter than the Asn618 one, the total protein volume is reduced.

Figure 3. Rph3A^{T437S} and Rph3A^{N605S} cause increase of total and GluN2A-dependent NMDA currents

(A) Representative traces of whole-cell responses to NMDA (50 mM) in Rph3A^{WT}, Rph3A^{T437S}**,** and Rph3A^{N605S} neurons show increase of I_{NMDA} amplitude when cells are transfected with mutant Rph3a. **(B)** Representative traces of whole-cell responses to NMDA (50 mM) and Ifenprodil (150 nM) in Rph3A^{WT}, Rph3A^{T437S} and Rph3A^{N605S} transfected neurons. Data show a 45% inhibition of I_{NMDA} in neurons overexpressing RFP-Rph3A^{WT} while neurons expressing RFP-Rph3A^{T437S} and Rph3A^{N605S} showed a 24% and 21% inhibition, respectively. **(C-D)** Representative bar graphs of total, GluN2A and GluN2B-mediated NMDA current **(C)** and charge **(D)**. RFP-Rph3A^{WT} (n=10); RFP-Rph3A^{T437S} (n=9); Rph3A^{N605S} (n=8). Two-way ANOVA.

Figure 4. GluN2A localization and availability and postsynaptic Calcium currents are impaired by *RPH3A* variants

(A) Confocal images of surface (magenta) and total GluN2A (green) staining in RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} transfected neurons (orange outline). Scale bar: 2 µm. **(B)** Bar graph representing mean ± SEM of percentage of RFP-Rph3A^{WT} mean GluN2A surface/total ratio, Kruskal-Wallis' test with Dunn's *post hoc* test. **p< 0.01. Dots represent single values. **(C)** Airyscan confocal images of GluN2A (green) and PSD-95 (magenta) staining in RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} transfected neurons (orange outline). Scale bar: 2 µm. **(D)** Bar graph representing mean ± SEM of percentage of GluN2A colocalizing with PSD-95, Ordinary One-way ANOVA with Dunnett's *post hoc* test. ***p<0.0001, **p<0.01. Dots represent single values. **(E)** Xph20-GCaMP7f fluorescent signals overtime (3 minutes) in DIV16 primary hippocampal neurons transfected at DIV8 with RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S}. **(F)** Bar graph representing mean ± SEM of the area under the curve (AUC) of spontaneous calcium transient events (evaluated as changes of Xph20-GCaMP7f fluorescence) at RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} and RFP-Rph3A^{T437S} dendritic spine over 3 minutes, Kruskal-Wallis' test with Dunn's *post hoc* test. **(G)** Violin plot representing the number of events per spine over 3 minutes, black lines represent mean, white dotted lines represent quartiles, Kruskal-Wallis' test with Dunn's *post hoc* test.

Figure 5. Evaluation of surface GluA1, spine density and morphology.

(A) Confocal images of surface (magenta) and total GluA1 (green) staining RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} transfected neurons (orange outline). Scale bar: 2 µm. **(B)** Bar graph representing mean ± SEM of percentage of RFP-Rph3A^{WT} mean of GluA1 surface/total ratio, Kruskal-Wallis' test with Dunn's *post hoc* test. Dots represent single values. **(C)** Confocal images of GFP filler in the dendritic spines of in RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S}

transfected neurons. Scale bar: 2 µm. **(B–F)**: Bar graph representation of mean ± SEM spine density **(D)**, spine length **(E)**, spine head width **(F)**, filopodia proportions **(G)** and spine types percentage **(H)** in RFP-Rph3A^{WT}, RFP-Rph3A^{N605S} or RFP-Rph3A^{T437S} transfected neurons. Ordinary One-way ANOVA with Dunnett's *post hoc* test (spine density, length, head width, mushroom, stubby), Kruskal-Wallis' test with Dunn's *post hoc* test (filopodia, thin). *$p < 0.05$. Dots represent single values.

Figure 1

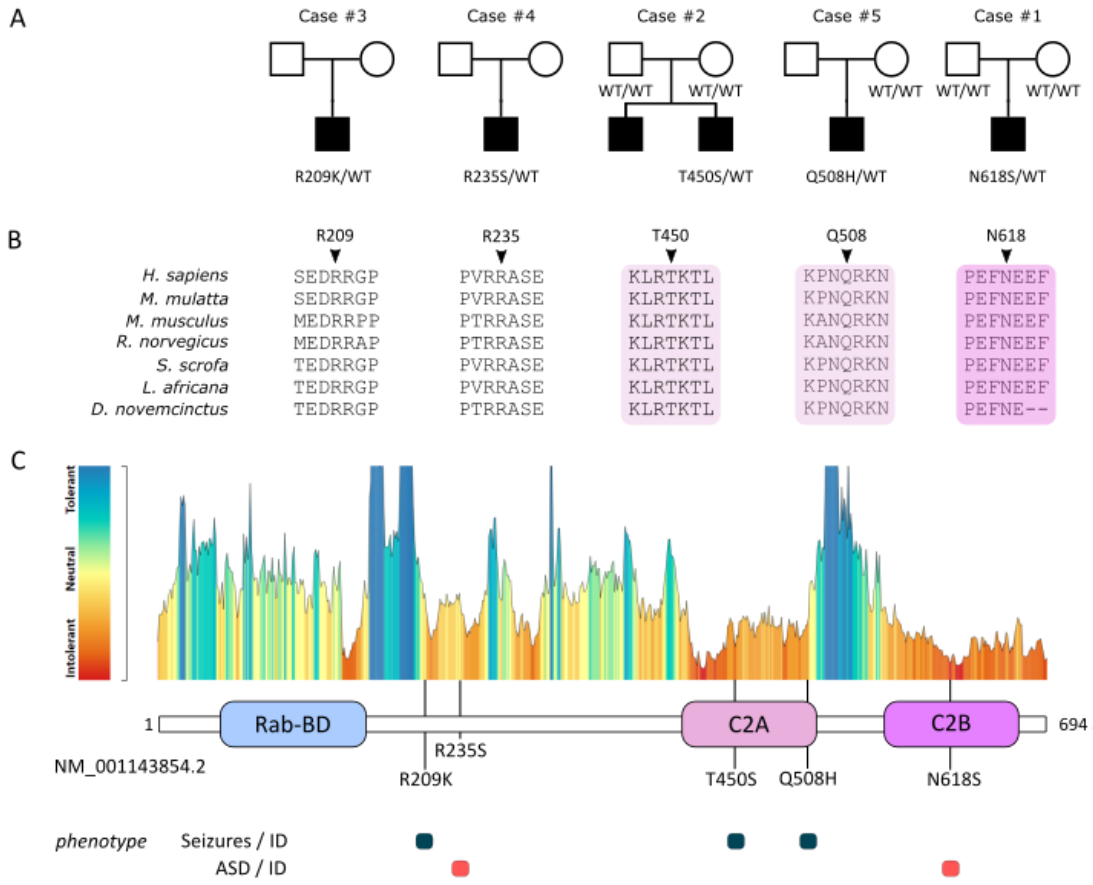


Figure 2

A

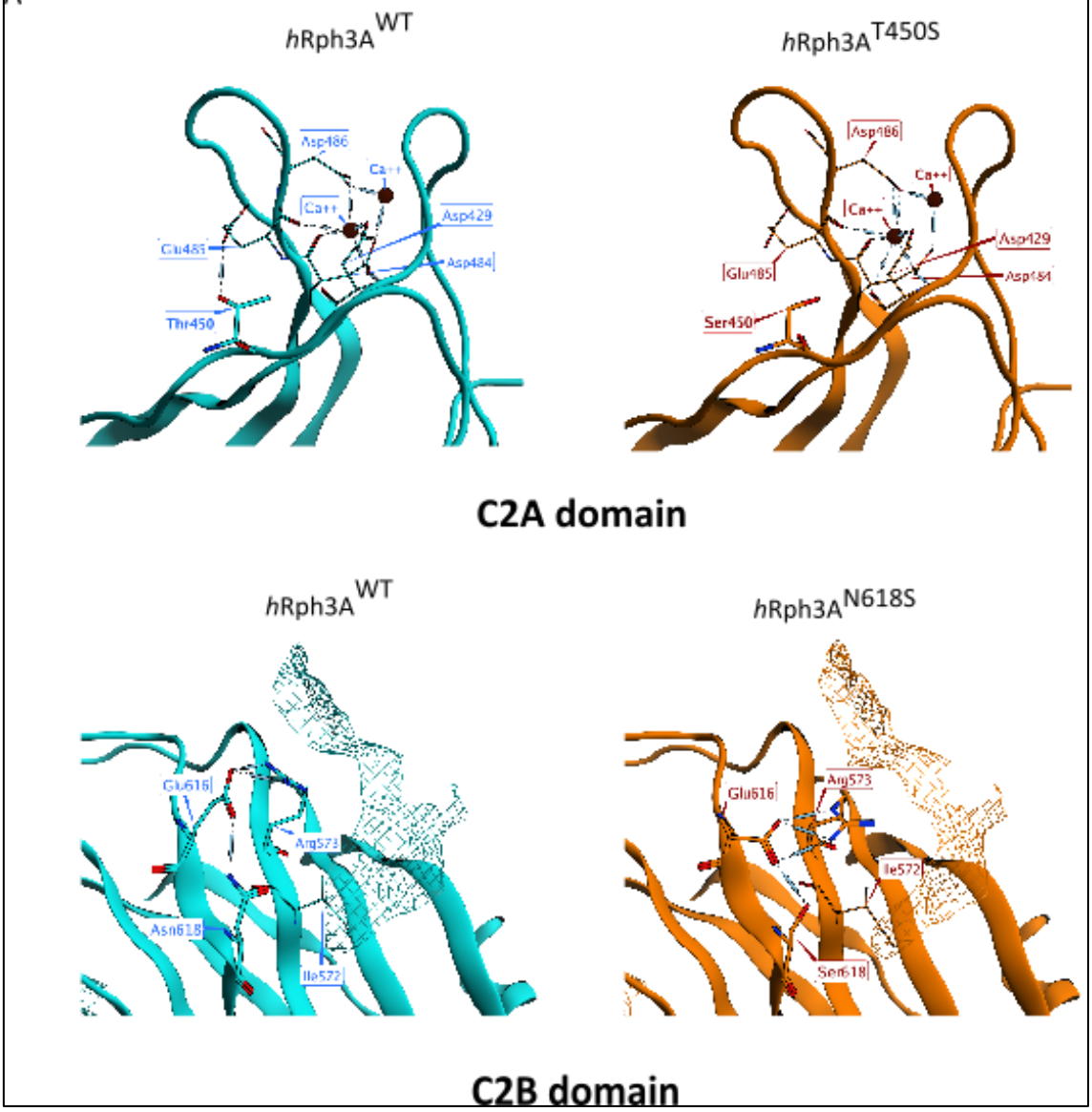

hRph3A^WT
hRph3A^T450S
Asp186
Ca++
Ca++
Asp429
Glu485
Asp484
Thr450
Asp186
Ca++
Ca++
Asp429
Glu485
Asp484
Ser450
C2A domain
hRph3A^WT
hRph3A^N618S
Glu616
Arg573
Asn618
Ile572
Glu616
Arg573
Ile572
Ser618
C2B domain

Figure 3

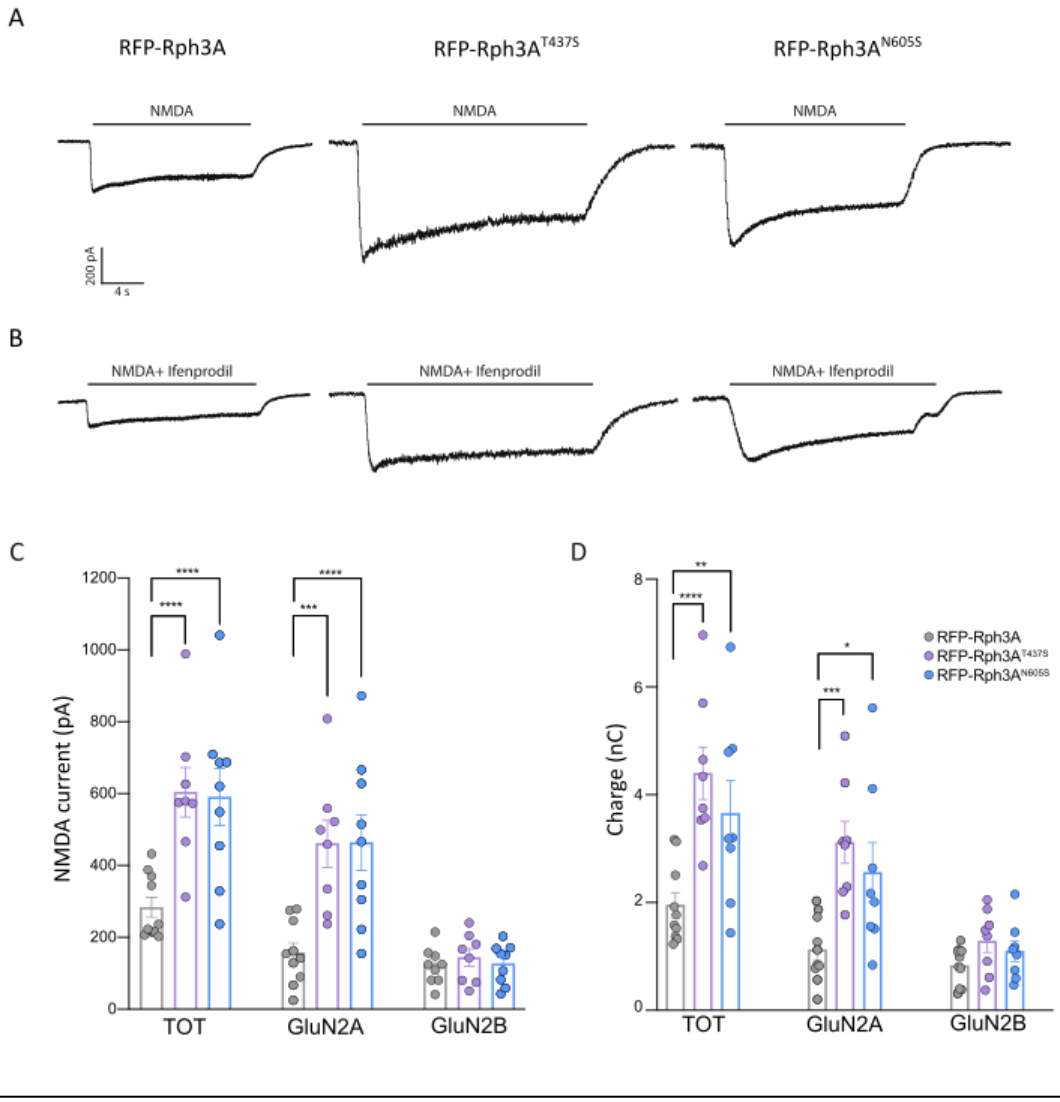

A
RFP-Rph3A
RFP-Rph3A^T437S
RFP-Rph3A^N605S
NMDA
NMDA
NMDA
200 pA
4 s
B
NMDA+ Ifenprodil
NMDA+ Ifenprodil
NMDA+ Ifenprodil
C
NMDA current (pA)
1200
1000
800
600
400
200
0
****
****
***
****
TOT
GluN2A
GluN2B
D
Charge (nC)
8
6
4
2
0
**
****
*
***
RFP-Rph3A
RFP-Rph3A^T437S
RFP-Rph3A^N605S
TOT
GluN2A
GluN2B

Figure 4

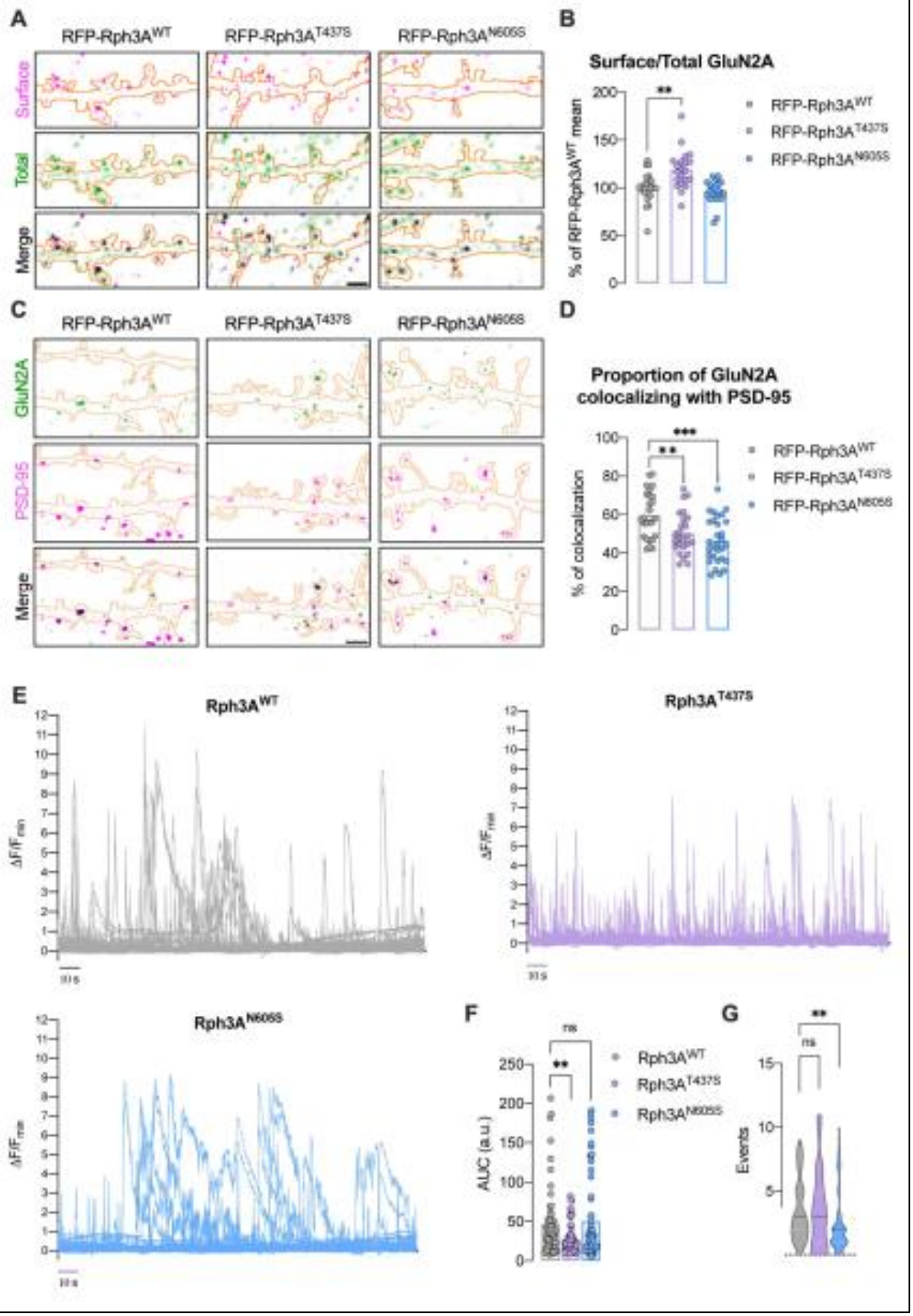
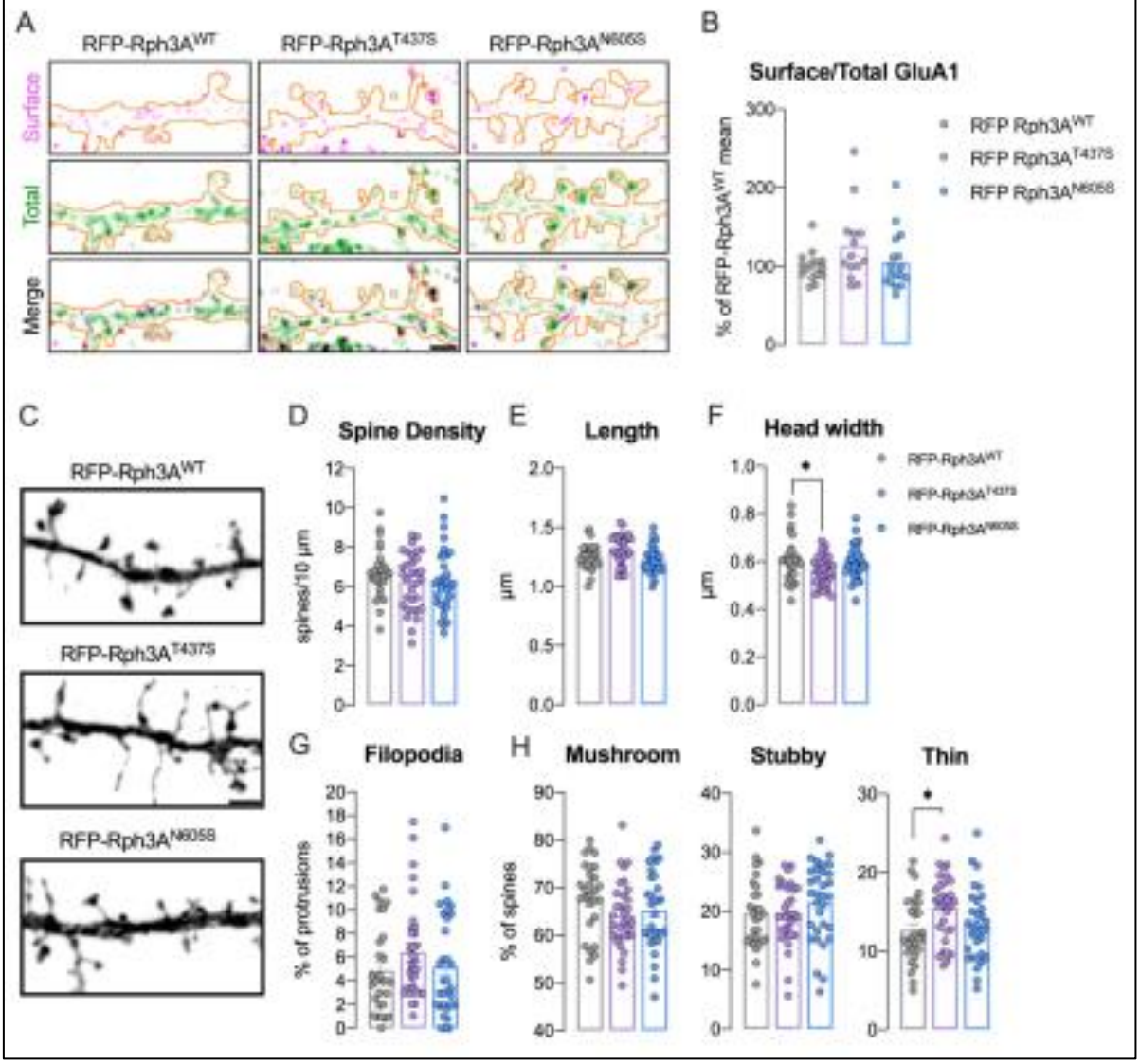

A
RFP-Rph3A^WT
RFP-Rph3A^T437S
RFP-Rph3A^N605S
Surface
Total
Merge
B
Surface/Total GluN2A
% of RFP-Rph3A^WT mean
RFP-Rph3A^WT
RFP-Rph3A^T437S
RFP-Rph3A^N605S
C
GluN2A
PSD-95
D
Proportion of GluN2A colocalizing with PSD-95
% of colocalization
E
Rph3A^WT
Rph3A^T437S
Rph3A^N605S
ΔF/F_min
F
AUC (a.u.)
Rph3A^WT
Rph3A^T437S
Rph3A^N605S
G
Events
ns

Figure 5


A
RFP-Rph3A^WT
RFP-Rph3A^T437S
RFP-Rph3A^N605S
Surface
Total
Merge
B
Surface/Total GluA1
% of RFP-Rph3A^WT mean
RFP Rph3A^WT
RFP Rph3A^T437S
RFP Rph3A^N605S
C
RFP-Rph3A^WT
RFP-Rph3A^T437S
RFP-Rph3A^N605S
D
Spine Density
spines/10 µm
E
Length
µm
F
Head width
µm
G
Filopodia
% of protrusions
H
Mushroom
% of spines
Stubby
Thin

Chapter 3: Novel mechanisms for known NDD genes

This section highlights the identification of plausible pathogenic molecular mechanism behind a known disorder, the Mental Retardation Autosomal Dominant 57 (MRD57), associated with variants in *TLK2* gene[22]: indeed, despite the many MRD57-patients described mechanisms behind the pathogenic mechanisms were still not investigated. The following paragraphs will attempt to shed light on this topic.

3.1 TLK2-related neurodevelopmental disorder: new insights on the pathological mechanisms

The Tousled-Like Kinases 1 and 2 (TLK1 and TLK2) are serine-threonine kinases involved in DNA replication and repair, transcription, cell cycle checkpoint recovery, chromatin maintenance and genomic stability[148–150]. Both kinases target the ASF1A and ASF1B histone H3/H4 chaperones and are regulated by DNA damage responsive checkpoint signalling[151]. Depletion of *Tlk1* and *Tlk2* in mice indicated that they are largely redundant, except for an essential role for *Tlk2* in placental development[152]. Null *Tlk2* mice, generated with a conditional allele to bypass the placental defect, showed no gross developmental defects except for a slight growth delay compared to controls, suggesting that TLK1 can compensate for loss of TLK2 function after embryonic development. Further work in human cancer cells supports extensive redundancy at the cellular level. Co-depletion of TLK1 and TLK2, however, causes replication stress, DNA damage and altered chromatin maintenance, particularly affecting telomeres and other repetitive genome elements, while mild effects were observed with depletion of either protein[153].

Mutations in *TLK2* were associated with Mental Retardation Autosomal Dominant (MRD57, MIM: 618050) by Reijnders *et al.*[22], who described 40 cases from 38 unrelated families. MRD57 is clinically characterized by autism spectrum disorder (ASD), intellectual disability (ID), behavioural problems, growth delay and facial dysmorphism, including blepharophimosis, telecanthus, prominent nasal bridge, broad nasal tip, thin vermilion of the upper lip and upslanting palpebral fissures. Other features shared by a subset of cases are gastrointestinal problems, seizures, skeletal malformations, and ocular problems.

In most reported cases, the disease is likely due to *TLK2* haploinsufficiency, as most of cases are heterozygous for loss-of-function (LoF) alleles, which is in line with the strong constraint against LoF variants in the gene (pLI = 1, GnomAD database). Reported missense variants cluster mainly in the C-terminal Protein Kinase Domain (PKD)[22,130,154], the core of TLK2 function. Four *TLK2* variants localized in this region have previously been analysed, showing a strong reduction of kinase activity on ASF1A *in vitro*[155]. TLK2 activation requires dimerization through an N-terminal

coiled coil motif, suggesting that inactive mutants could have a dominant negative effect[155]. Thus, based on the available data, the predominant pathogenic mechanism of *TLK2* mutations appears to be a reduction in its overall activity. Of note, recent work suggested a possible autosomal recessive phenotype in a proband affected by severe neurodevelopmental disease with a homozygous missense variant (c.163A>G; p.(Lys55Glu))[156]. This variant is localized outside the TLK2 PKD and coiled-coil motifs, and may be hypomorphic, since the carrier parents are not affected.

During the screening of a large survey of patients with ID by array-CGH and trio-ES, we identified three new cases with *TLK2*-mutations. Two subjects were heterozygous for a *de novo* 39-kb deletion encompassing *TLK2,* and a *de novo* c.1652A>G; p.(Asp551Gly) missense change. The third case was familial, with a nonsense variant (c.1423G>T; p.(Glu475Ter)) occurring in three affected siblings and their affected mother. Here, we report their clinical description, confirming and expanding the disease phenotype. We also provide data documenting an altered chromatin state in patient-derived fibroblasts and lymphoblastoid cells (LCLs), consistent with a defect in the regulation of histone chaperones. We also characterized the activity and proximal interactomes of the p.(Asp551Gly) variant, as well as another missense variant (p.(Ser617Leu)), reported in a several published studies[26,157,158]. Both variants exhibited impaired kinase activity and TLK2 proximal interactomes were enriched with proteins previously implicated in ID and ASD, suggesting connections to a larger chromatin maintenance network. Due to the important involvement of TLK2 chromatin remodelling, we recently expanded our cohort of MRD57-cases, and started to analyse the methylation profile on patients' DNA[96]. Preliminary results, show a detectable episignature[38,39] in the analysed patients, opening novel perspectives for functional studies and clinical approaches.

Materials and methods

Whole exome sequencing, prioritization, and variant calling

DNA was extracted from total blood using the ReliaPrep Blood gDNA Miniprep kit (Promega, Madison, WY, USA) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientifics, Waltham, MA, USA). Informed consent was obtained from participating families and the study protocol was approved by our internal Ethics Committee, according to the Declaration of Helsinki.

Array-CGH was performed using a 60K whole-genome oligonucleotide microarray (Agilent Technologies, Santa Clara, California, USA). Family 1 and 2 were enrolled in the Autism Sequencing Consortium (ASC) project and their gDNA samples were sequenced at the Broad

Institute on Illumina HiSeq sequencers as previously described[6,26]; variant calling was performed using targeted bioinformatic pipelines adapted for different pattern of inheritance. Identified variants were confirmed by Sanger sequencing using standard conditions and the primers in table S1. Additional information is provided in Supplementary materials and methods.

All variants are referred to GRCh37 annotation and to NM_001284333.2, in line with the previously published TLK2 structure[155]. For homogeneity with the clinical work from Reijnders *et al*.[22], we specified variants also in NM_006852.6 in Supplementary table S2.

In silico prediction of missense variants impact and splicing analysis

Variants were analysed with the VarSome tool[159] as a starting point for further analysis. This allowed evaluation of at least 15 *in silico* predictors simultaneously. Variants frequencies were evaluated using Genome Aggregation Database (GnomAD) Browser version 2.1.1. Impaired splicing was predicted using Human Splicing Finder (HSF) version 3.1[160] and experimentally verified as described in Supplementary materials and methods.

Cell cultures

Peripheral blood mononuclear cells (PBMCs) were isolated from whole blood using Histopaque®-1077 (Sigma-Aldrich) and subsequently immortalized with Epstein-Barr Virus (EBV) and cultured in RPMI medium (GIBCO, Thermo Fisher Scientific) supplemented with 10% Fetal Bovine Serum (FBS) (GIBCO), 1% Pen-Strep and 1% L-Glutamine. Primary fibroblasts were isolated from human skin biopsies after overnight incubation in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich) supplemented with 10% FBS and 160 µg/ml collagenase. Fibroblasts were maintained in DMEM supplemented with 10% FBS, 1% Pen-Strep and 1 mM sodium pyruvate (Thermo Fisher Scientific) at 37°C, 5% CO_2. AD-293 cells (Stratagene) were grown in DMEM supplemented with 10% FBS (Sigma-Aldrich), 50 U/mL penicillin and 50 µg/mL streptomycin (Thermo Fisher Scientific) at 37ºC in a 5% CO_2 incubator.

RNA isolation and quantitative real time PCR

Total RNA was extracted from fibroblasts and LCLs using the Direct-Zol RNA MiniPrep system (Zymo Research, Irvine, CA, USA) and complementary DNA (cDNA) was generated using the M-MLV Reverse Transcriptase kit (Invitrogen, Thermo Fisher Scientific). The expression level of *TLK2* was measured using the Universal Probe Library system (Roche Diagnostics, Risch-Rotkreuz, Switzerland) with primers and probe in table S1 and *HMBS* and *GAPDH* as reference gene (assays numbers Hs00609297_m1 and Hs00266705_g1; Applied Biosystems, Thermo Fisher Scientific). Assays were carried out in triplicate on an ABI 7500 real-time PCR instrument using the ABI 2X TaqMan™ Universal PCR Master Mix II, according to the manufacturer's protocol (Thermo Fisher Scientific). For each experiment, biological triplicates with at least two

technical replicates were performed. Data were analysed with Prism-GraphPad Software performing unpaired t-test with Welch's correction. p-values are indicated as follows: ns= P value > 0.05 ; *= P value ≤ 0.05; ** P value ≤ 0.01; ***= P value ≤ 0.001; ****= P value ≤ 0.0001.

Single-cell gel electrophoresis

Samples were prepared according to the alkaline single-cell gel electrophoresis (SCGE) assay method, as previously described[28] and are briefly summarized in Supplementary material and methods, together with experimental details.

Site directed mutagenesis and in vitro kinase assays from cell lysates

In vitro kinase assays were performed as previously described[155] with minor modifications. Full methods provided in the Supplementary materials and methods and primers in table S3.

Western blotting

For affinity purification (AP), 40 µg of input protein and 20 µL of Strep-AP elution, with 6X SDS (0.2% Bromophenol blue and *β*-mercaptoethanol), were separated by SDS-PAGE and transferred to nitrocellulose membranes (0.2 or 0.45 µm pore, Amersham Protran; Sigma-Aldrich). For detection of Streptavidin, PVDF membranes (0.45 µm pore, Immobilon-P, Merck) were used. Membranes were blocked and antibodies prepared in 5% non-fat milk in PBS-T, except for CHD7 and CHD8, where 5% BSA in PBS-T was used. Primary antibodies were detected with the appropriate secondary antibodies conjugated to Horseradish peroxidase (HRP) (table S4) and visualized by ECL-Plus (GE Healthcare).

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

BioID-MS and BioID-Westerns were performed essentially as described in Silva *et al.* [161] with some modifications. Full methods are provided in the Supplementary materials and methods and data is available in the PRIDE database with accession number PXD019450.

Methylation data generation and analysis

Peripheral blood DNA was extracted using standard techniques. Bisulfite conversion was performed with 500 ng of genomic DNA using the Zymo EZ-96 DNA Methylation Kit (D5004), and bisulfite-converted DNA was used as input to the Illumina Infinium HumanMethylation450 (450K array) or MethylationEPIC BeadChip array (EPIC array). Array data were generated according to the manufacturer's protocol. Sample quality control was performed using the R minfi package version 1.35.2.34. The data analysis pipeline was performed as previously described[39].

Results

Identification of novel MRD57 patients

The identification of *TLK2* as a risk gene for intellectual disability[162], prompted us to re-evaluate the genomic information available for a large cohort of patients affected by ASD and/or ID who had previously been analysed by trio-ES. This included our in-house cohort (2,250 total samples, 736 affected), as well as a publicly available one from the Autism Sequencing Consortium (35,584 total samples, 11,986 affected)[26]. No further likely pathogenic variants or variant of unknown significance (VUS) in neurodevelopment-related genes[149] were identified. A search in the DECIPHER database[163] for novel *TLK2* deletions was also performed. We found six cases in three independent families with detrimental variants in the *TLK2* gene (figure 1A-F).

In family 1, the proband carried a *de novo* missense change c.1652A>G; p.(Asp551Gly) not reported in gnomAD and predicted to be damaging by multiple *in silico* predictors (table S5). This variant localizes in a highly Constrained Coding Region (CCR) (>93th percentile) within the PKD[164]. In family 2, three affected siblings carried a premature stop variant (c.1423G>T; p.(Glu475Ter)) inherited from their affected mother (case 2). This variant was not reported in gnomAD and was classified as pathogenic using the American College of Medical Genetics (ACMG) criteria[156,162]. A sixth case (family 3) was identified in the DECIPHER database: a female carried a likely pathogenic (class 4)[156,162] *de novo* del(17)(q23.2), whose minimum size was 39 kb (NC_000017.10:g.(60683462-60722398)del). The deletion encompassed *TLK2*, *MRC2* and *MARCH10* genes. Neither the *MRC2* or *MARCH10* genes are predicted to be haploinsufficient (GnomAD *MRC2* pLi=0.67; *MARCH10* pLi=0). *MRC2* encodes the Mannose Receptor C Type 2 and has a role in the turnover of collagens in the cytomembrane and extracellular matrix[154]. Its upregulation has been linked to tumorigenic activity, and a role in breast cancer and hepatocellular carcinoma has been suggested[130,155] but, to our knowledge, there is no supporting literature implicating *MRC2* in neurodevelopmental disorders. The literature available for *MARCH10* is limited but the gene is expressed almost exclusively in testis (https://gtexportal.org/home/gene/MARCH10) and was suggested to play a role in spermiogenesis[26]. Therefore, we think it is unlikely that haploinsufficiency in *MRC2* or *MARCH10* contribute to the patient's phenotype.

Our patients (3 males and 3 females) ranged from 3 to 47 years of age, all of them were of Caucasian ethnicity. A broad range of behavioural disorders was present, including ASD (1/6), attention deficit hyperactivity disorder (ADHD) (4/6), anxiety (3/6), short attention span (2/6) and obsessive-compulsive behaviour (2/6). ID was reported for 4 patients in the borderline (IQ 70-85) or low (IQ≤70) range. Patient 5 was too young for formal assessment of a neurodevelopmental phenotype, but a global developmental delay was reported. No formal

evaluation for intellectual disability was provided for case 6 but difficulties in memory and transcription were reported. All affected members from family 2 had microcephaly. Dysmorphic facial features were observed in all patients and included upslanting palpebral fissures, wide nose, low hanging columella, smooth philtrum, prognathism, pointed chin and hypertelorism (figure 1G). Minor skeletal anomalies of the hands and feet were reported for four patients (figure 1H). Interestingly, some of the features we observed in a portion of our patients had not been described in previously reported cases, expanding the clinical phenotype. Among them, there were neurodevelopmental (difficulties in reading, writing, memory, and transcription and pavor nocturnus), dysmorphic (prognathism, bifid nasal tip, low hanging columella, absent ear lobe, low-implant auricle, synophrys and downturned corners of mouth) and skeletal anomalies (postaxial polydactyly of left foot, overriding second toe, tapering fingers, short hands with short distal phalanx). A summary of the clinical phenotypes is reported in table S6 and a description is provided in the Supplemental Data, while a comparison with the available literature is provided in table S7.

TLK2 haploinsufficiency disrupts proper chromatin compaction

Using quantitative real time polymerase chain reaction (qRT-PCR), we analysed *TLK2* mRNA expression in LCLs from case 1 and fibroblasts from case 6. We found an approximately 50% reduction in *TLK2* mRNA expression in both cases (figure 2A and S1A). This was unexpected in case 1, who carried the c.1652A>G; p.(Asp551Gly) missense change. Variants encoding missense changes may induce unstable mRNA secondary structures leading to degradation of the altered allele[130,154]. To verify if both alleles were expressed, we amplified and sequenced cDNA between exons 18 and 23. We found that both wild type and mutant allele were present (figure S1B), suggesting that the allele encoding the p.(Asp551Gly) variant was not degraded. Moreover, no differences were observed in band sizes from WT and p.(Asp551Gly) cDNAs (figure S1C), excluding that p.(Asp551Gly) led to the production of aberrantly spliced transcripts[165,166]. Western Blot analysis on LCLs from case 1 compared to controls, uncovered a significant increase in TLK2 protein expression (figure S1D), motivating us to perform a deeper characterization of the impact of p.(Asp551Gly) on TLK2 activity.

Based on the role of TLK2 in controlling chromatin remodelling[149], we investigated possible changes in chromatin compaction, performing the single-cell gel electrophoresis (SCGE) assay with LCLs derived from case 1. Relatively short electrophoresis time was already sufficient to unmask slight differences between control and mutant cells. Longer run times allowed DNA loops to stretch under the electric field and revealed a significantly more relaxed state of

nucleoids in LCLs from the affected subject compared to control cells (figure 2B-C). Differences were quantified as "tail moment" values, which are defined as the product of the tail length and the percentage of DNA in the tail. LCLs harbouring the p.(Asp551Gly) variant also exhibited a higher sensitivity to γ-ray irradiation, documenting increased susceptibility to DNA damage (i.e., single and double strand breaks), which is in line with a more relaxed state of chromatin (figures 2D, E). Of note, a defective heterochromatin state was also observed in fibroblasts derived from subject 6 carrying the 17q23.2 deletion encompassing *TLK2* (figure S2), suggesting a similar effect between the 17q23.2 deletion and the p.(Asp551Gly) variant. Overall, these findings demonstrated that TLK2 haploinsufficiency disrupts proper chromatin organization and confers susceptibility to DNA damage.

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

All the *TLK2* missense mutations in the PKD from MRD57 analysed to date exhibited decreased kinase activity, including H493R, H518R, R720A and D629N[155](Figure S4B). Given that the high conservation of the mutated residues in the PKD suggested reduced kinase activity, we examined the potential structural impact of the missense p.(Asp551Gly) (hereafter indicated as D551G) variant identified in our survey, as well as the p.(Ser617Leu) (hereafter indicated as S617L) variant, found in a patient with ASD already described in whole exome sequencing works[26,157,158]. The patient was also reported in the ASC exome analysis browser, a freely accessible portal containing de novo variants identified in more than six thousand affected probands enrolled in ASC WES projects (https://asc.broadinstitute.org/). Residue S617 is located one residue after the Asp-Phe-Gly (DGF) motif[167,168], that together with the activation and catalytic loop, constitute the kinase core[155]. We previously identified S617 as a TLK2 auto-phosphorylation site and demonstrated that alterations of this site can significantly enhance (S617A) or reduce (S617D) TLK2 kinase activity[155]. Therefore, it remained an open question how S617L, associated with autism spectrum disorder, would influence the kinase activity of TLK2.

We modelled both mutations using the crystal structure of the TLK2 PKD[155]. TLK2-D551G was predicted to weaken hydrogen bonds with the subsequent helix and S617L introduces a hydrophobic residue in place of the auto-phosphorylation site in the activation loop (figure 3A-B).

To determine if these mutations affected kinase activity, we analysed TLK2 activity using *in vitro* kinase assays. Ectopically expressed, Strep-FLAG tagged TLK2, TLK2-KD (kinase dead, D592V), TLK2-D551G and TLK2-S617L were affinity purified from AD-293 cells using an N-terminal Strep tag and incubated with purified substrate (ASF1A) for kinase assays. Both mutations led to a

notable reduction in substrate modification (figures 3C-D). Quantification of multiple experiments demonstrated that TLK2-S617L severely impaired kinase activity, comparable to the TLK2-KD protein, while TLK2-D551G was more mildly impaired and showed slightly higher autophosphorylation levels than TLK2-WT in some experiments (figure 3D).

In previous work, we noted that loss of the coiled-coil domains of TLK2 led to perinuclear accumulation[155]. To determine if the TLK2-D551G and TLK2-S617L missense mutations altered TLK2 localization, we transfected FLAG-tagged mutants in AD-293 cells and performed immunofluorescence (IF) microscopy. Lamin A was used as an inner nuclear membrane marker and nuclear DNA was stained with DAPI. TLK2-WT showed diffuse nuclear localization in transfected cells, as observed previously (figures 3E, S3A). In contrast, TLK2-D551G and TLK2-S617L showed perinuclear localization to different extents (figures 3E-F, S2A). This was particularly prominent for the TLK2-S617L mutant, where 75% of transfected cells showed a perinuclear localization of TLK2 (figures 3E-F).

The proximal interactome of TLK2 is altered by missense mutations

TLK2 is involved in many biological processes and few clear substrates aside from ASF1A and ASF1B have been well characterized[149]. As kinases often bind to substrates with low affinity, we previously used an unbiased proximity biotinylation assay coupled to mass spectrometry (BioID-MS), as it does not require high-affinity interactions that can withstand purification procedures[152,169]. We used this approach to further characterize the cellular environment of TLK2 and determine if the missense mutations influenced its interactome.

BirA-tagged *TLK2-WT*, *TLK2-D551G* and *TLK2-S617L* were expressed in AD-293 cells by transient transfection (figure 4A). Network clustering of results from wild type TLK2 (TLK2-WT) identified the known TLK2 substrates ASF1A, ASF1B and TLK1, as well as the DYNLL1/2 (LC8) proteins that we previously validated[152,170]. The proximal interactome grouped into five functional clusters consistent with the known functions of TLK2, including RNA processing and splicing, transcriptional regulation, chromatin binding or remodelling, DNA repair and histone chaperones (figure 4B and table S8). We cross referenced the proximal interactome with a recent large-scale analysis of iPOND-MS data that identified TLK1 and TLK2 as high confidence interactors with nascent DNA at active replication forks[171]. Several of the TLK2-WT hits, including RAD50, BRD4, CHD8, ASF1B, SCML2 and NACC1 were also found at active forks with high confidence in iPOND studies. We next compared our TLK2-WT proximal interactome to the SFARI (Simon's Foundation Autism Research Initiative) and DECIPHER databases and identified 8 proteins: CHD7, CHD8, NACC1, CCNK, JMJD1C, RAD50, MSANTD2, and YEATS2. These data

suggest potential functional links between TLK2, and a number of proteins involved in neurodevelopmental disorders with overlapping pathologies. Details about SFARI and OMIM classifications are provided in Supplementary table S9.

In parallel to TLK2-WT, we performed BioID analysis with TLK2-D551G and TLK2-S617L. Both mutants accumulated to higher levels than TLK2-WT, consistent with what we previously observed with other inactive TLK2 mutants[155]. Both mutants caused numerous alterations in the proximal interactions compared to TLK2-WT (figures 4C, D and 5). This included a reduction in several replication fork and ASD related proteins, including RAD50 and YEATS2 with both mutants, and JMJD1C, BRD4, CCNK, NACC1, MSANTD2 and CHD8 with TLK2-S617L (figures 4C, D and 5). In addition, some proteins, including ZNF148, NFIA, NFIX and PAPOLG, that are all in the SFARI database, were significantly enriched with both mutants, but not TLK2-WT (figures 4C, D and 5).

The CHD7 and CHD8 chromodomain helicases were of particular interest to us because they are mutated respectively in CHARGE syndrome[172] (OMIM 214800) and susceptibility to autism[173] (OMIM 615032) and implicated in modulating chromatin structure[174,175]. In addition, CHD8 has been localized to active sites of DNA replication, like both TLK1 and TLK2[171]. CHD8 spectra were detected at similar levels between TLK2-WT and TLK2-D551G, that retains some activity, but were reduced with the kinase dead TLK2-S617L (figures 4C, D and 5). In contrast, CHD7 spectra were highest in cells expressing TLK2-D551G. We performed BioID-Westerns to validate these proximal interactions and their relative differences. Expression of the control N-FLAG-BirA and biotin supplementation led to no detectable CHD7 or CHD8 detected in Western blots of Strep-affinity purified protein lysates (figure 6A). In contrast, both CHD7 and CHD8 were clearly validated with all TLK2 alleles, although CHD8 levels were lowest with S617L and CHD7 highest in TLK2-D551G, consistent with the MS data. As expected from the BioID-MS, TLK1 co-purified with all TLK2 alleles to a similar degree. It was also notable that the substrates, ASF1A and ASF1B, were highest with D551G and similar between TLK2-WT and TLK2-S617L, despite the much higher level of TLK2-S617L expression. These results indicate the missense mutations have differential effects on both proximal interactions and TLK-ASF1 interactions.

MRD57 is associated with a detectable episignature in patients' DNA (preliminary results)

The evidence of an involvement of TLK2 in chromatin remodeling prompted us to evaluate eventual changes in DNA methylation profiles in MRD57, as for *CAPRIN1* patients (*see* Chapter 2). The evaluation of eight patients (the six described above plus three additional cases, see Table 1) with overlapping clinical phenotypes, revealed preliminary supporting data about the

presence of a recognizable methylation profile (Fig. 7). The analysis of six additional cases (Table 1) is ongoing.

Among the cases, we analyzed a patient with a double genetic cause of genetic disorder: the *de novo* variant c.1480C>T; p.(Arg494Trp) in *TLK2* and the *de novo* variant c.11182C>T; p.(Arg3728Ter) in *DYNC1H1* (NM_001376.5), the last associated with MRD13 (OMIM #614563). The clinics associated with *TLK2* and *DYNC1H1* are very similar, and the variants are respectively classified as likely pathogenic and pathogenic[34]. It is interesting to note that this case did not cluster with other *TLK2* cases, suggesting a possible regulation of DNA methylation performed by both mutated genes, resulting in a methylation profile different from the other ones. Additionally, BioID experiments suggest an interaction between TLK2 and DYNC1H1, suggesting the possible presence of complex mechanisms of co-regulation (Table S9). Further experiments are needed to clarify this point and the impact of each variant on DNA methylation.

Discussion

To our knowledge, at least 13 *TLK2* missense variants have been reported in MRD57 cases [162,176,177], with mutations clustering in the PKD. Previously described patients showed ID in different range of severity, ASD, language delay, motor delay, gastro-intestinal issues, and dysmorphic facies as prevalent features[22,154]. In our cases, we confirmed a high incidence of ID (4/4), language delay (4/6) and dysmorphic facial features (6/6). On the contrary, comparing to the available clinical literature, we observed a under representation of gastro-intestinal problems (0/6) and ASD (1/6). We also observed additional features, including difficulties in reading and writing (3/6) or in memory and transcription (3/6) and skeletal anomalies of the hands (1/6) and the feet (1/6).

Functional characterization has been reported only for four variants in the PKD, (p.(His493Arg), p.(His518Arg), p.(Asp629Asn), p.(Arg720Ala)) and has shown at least a 50% reduction in enzymatic activity compared to wild type protein[170]. Our data further expands the characterization of MRD57 missense mutations and reinforces the prevailing hypothesis that the majority of these impair TLK2 kinase activity. Both TLK2-D551G and TLK2-S617L showed profoundly impaired kinase activity, as well as altered subcellular localization, the significance of which remains unclear (figure 3, S3 and S4). As also suggested by *in silico* tools (table S5), TLK2-S617L caused a more severe impairment of protein activity and localization compared to TLK2-D551G mutant. It is interesting to note, that case carrying p.(Asp617Leu) was characterized by a mild-severe intellectual disability, with a IQ (verbal IQ 63, non-verbal IQ 74) that was lower than the one observed in case carrying p.(Asp551Gly) variant and comparable to the one

observed in patients carrying p.(Glu475Ter) variant (table S6). Basing on the limited available clinical information, we could therefore suggest that the higher impairment of the protein activity mediated by p.(Ser617Leu) variant could be reflected on the clinical side by a more severe phenotype.

Both mutants overexpressed in AD-293 cells showed higher expression compared to WT. Accordingly, LCLs carrying p.(Asp551Gly) variant showed significantly higher TLK2 protein levels compared to control. Surprisingly, the higher protein expression was counterbalanced by drastically reduced mRNA levels, suggesting a potential negative feedback control, further causing reduced *TLK2* total mRNA levels.

Despite the severe effects on kinase activity caused by the heterozygous missense mutations identified in MRD57 patients to date, monoallelic *Tlk2* loss did not cause overt phenotypes in mice, but neurodevelopment and behaviour were not assessed in these animals[152]. Given that TLK2 dimerizes with both TLK1 and TLK2 and this is important for its activity, it is also likely that kinase impaired mutants exert some dominant negative effects that contribute to an overall phenotype that is more severe than haploinsufficiency[178]. The recent identification of multiple MRD57 cases with *TLK2* haploinsufficiency suggests that an in-depth evaluation of neurodevelopment is needed in mice with reduced *Tlk2* levels to determine if they represent a model of MRD57. The placental issues observed in mice with homozygous deletion of *Tlk2* were not identified in *Tlk2* heterozygous mice, but human gestation is considerably longer, and more subtle placental issues could be present. Recent work showed that total TLK depletion leads to an innate immune secretory response in cancer cells and mice[149,153,179]. Maternal immune activation (MIA) has been implicated in ASD and associated with placental defects in mice, suggesting that impaired chromatin maintenance and epigenetic dysregulation could potentially underlie the pathological effects of TLK2 haploinsufficiency[149,180,181]. This is consistent with the increased chromatin accessibility reported here in TLK2-D551G patient cells (figure 2B-E), as well as the strong enrichment of chromatin proteins in the TLK2 interactome in the SFARI and DECIPHER databases.

Many genes encoding proteins involved in chromatin remodelling are associated with neurodevelopmental disorders. TLK2, as well as the missense mutants we tested, showed proximal interactions with many of them, including CHD8 and CHD7, that are mutated in ASD and CHARGE syndrome. In addition to these proteins, both TLK2 missense mutants showed altered interactions with additional proteins implicated in neurodevelopment. This included RAD50, a part of the MRE11-RAD50-NBS1 DNA repair complex that localizes to replication forks and plays a key role in DNA-double-strand break repair[182]. *RAD50* mutations, present in the

DECIPHER database, underlie Nijmegen breakage syndrome-like disorder (OMIM 613078). This condition is characterized by microcephaly, which is also commonly observed in many patients with *TLK2* variants. Further, RAD50 proximal interactions were reduced with both missense mutants, potentially suggesting reduced localization to replication forks (figure 5)[183]. Similarly, YEATS2, a chromatin reader component, is suggested as an ASD-associated gene by de novo genetic risk analysis and GWAS (SFARI database criteria 3.1, suggestive evidence)[11,12], was linked to epilepsy and was enriched with TLK2-WT compared to either missense mutant[184–186]. In contrast, ZNF148 and PAPOLG, that are also associated with neurodevelopmental disorders, were strongly enriched with both missense mutants, and detected at very low levels with TLK2-WT, while other SFARI genes, including *BRD4, JMJD1C, MSANTD2, CCNK* and *NACC1* were reduced specifically with the less active TLK2-S617L variant. In future work, it will be of interest to examine the potential functional relevance of these interactions to determine if their alterations underlie the altered chromatin state, we observed in LCLs from case 1 or other phenotypes associated with TLK2 loss of function. This approach may detect new candidate genes involved in neurodevelopmental disorders or help us understand the involvement of this network of SFARI genes in isolated or syndromic ASD and ID.

The knowledge of altered protein interactomes is important to understand the molecular impact of disease mutations and could be helpful in identifying pharmacological treatments to mitigate more severe phenotypes, such as epileptic seizures. It is attractive to imagine the possibility of repurposing drugs able to modulate the functions of some genes to influence their impact on disease pathology[187].

In conclusion, we provided the clinical description of six new cases carrying likely pathogenic and pathogenic *TLK2* variants and we presented new insights into the impact of *TLK2* missense variants, observing impairment in kinase activity, localization, and interaction (Figure S4). Our work offers a deep characterization of two missense variants localized in a key domain of the TLK2 protein, where most mutations related to MRD57-disorder occur, providing new insights into the potential role TLK2 in neurodevelopmental disorders.

Figures and tables

Table 1. *TLK2*-patients included in EpiSign analysis

ID	Sex	Age	Variant (hg19, NM_006852.6)	HPO terms
MS6761	M	10	c.31C>T; p.(Arg11Ter)	HP:0001249
MS6762	F	16	del17q23.2	HP:0002354; HP:0025356; HP:0001155; HP:0000271
MS6763	M	5	c.1586°>G; p.(Asp529Gly)	HP:0000717; HP:0001256; HP:00012433; HP:0007018; HP:0000722; HP:0001155

MS6764	F	21	c.1357G>T; p.(Glu453Ter)	HP:0001256; HP:00012433; HP:0000722
MS6765	M	11	c.1357G>T; p.(Glu453Ter)	HP:0002187; HP:00012433; HP:0007018; HP:0001760
MS6766	F	43	c.1357G>T; p.(Glu453Ter)	HP:0002187; HP:0012433; HP:0007018; HP:0001760
MS6767	M	2	c.1357G>T; p.(Glu453Ter)	HP:0012433; HP:0007018
MS6768	F	14	c.1357G>T; p.(Glu453Ter)	HP:0001249; HP:0001263; HP:0001273; HP:0002650; HP:000118; HP:0002353
2022D00243_D01	F	2	c.1357G>T; p.(Glu453Ter)	HP:0000252; HP:0000271; HP:0012758; HP:0002463; HP:0000708
1898	F	1.1	c.989C>A; p.(Ser330Ter)	HP:0011318; HP:0000506; HP:0002553; HP:0000581
21NR00412	M	20	c.1794C>T; p.(Ser595Leu)	HP:0001256; HP:0000717; HP:0001007; HP:0000276; HP:0003189; HP:0000160; HP:0000233
D/3564/19-1	M	9.8	c.2034del; p.(Glu679fs)	HP:0001256; HP:0000271; HP:0000736; HP:0001763; HP:0000708
25867	M	17	c.1579G>A; p.(Gly527Arg)	HP:0001249; HP:0000717
dl_at	F	5.6	c.1889G>C; p.(Gly630Ala)	HP:0001249

Figure 1. Facial features and skeletal anomalies of individuals with *TLK2* variants

(A-C) Pedigrees of family 1, 2 and 3. Cases from family 1 and 3 carried respectively a heterozygous *de novo* missense variant (c.1652A>G; p.(Asp551Gly)) and a heterozygous *de novo* deletion encompassing *TLK2* gene. Cases from family 2 shared a heterozygous premature stop variant (c.1423G>T; p.(Glu475Ter)) inherited from an affected mother. Analysis of maternal grandparent genotype were not possible, but familial clinical history did not suggest a possible MRD57-like phenotype for them. wt= wild type at variant position. **(D-E)** Sanger validation of variants identified in family 1 and 2. Validation was performed both on gDNA from affected cases and from their unaffected relatives. **(F)** 17q23.2 deletion (minimum size 39 kb, chr17-60683462-60722398) identified in case 6 from family 3. The 39 kb deletion encompassed *TLK2* and *MRC2* genes. **(G)** Frontal and lateral face photographs of our cases, showing overlapping facial dysmorphisms. Most frequently reported features were upward slanted palpebral fissures, broad nasal tip, thin lips, low hanging columella, prognathism, wide spaced eyes and downturned corners of the mouth. **(H)** Details of reported skeletal anomalies at the level of hands and feet. Upper left panel: tapered hands fingers from case 1; upper right panels: short hands with short distal phalanx from case 6; bottom left panel: postaxial polydactyly of left foot observed in case 3; bottom right panel: overriding second toe in case 4.

Figure 2. TLK2-Asp551Gly affects chromatin density and confers susceptibility to DNA damage.

(A) *TLK2* mRNA levels in fibroblasts from case 6 and in lymphoblastoid cell line derived from case 1. *TLK2* expression was significantly reduced both in LCL carrying p.(Asp551Gly) variant and in fibroblasts carrying the 17q23.2 deletion. All experiments were performed at least in triplicate. UPL probe #72 and primers indicated in Supplementary material and methods were used; *HMBS*

mRNA expression was used as reference. Statistical analysis was performed using t-test with Welch's correction; ****= P value ≤ 0.0001. Numbers at the top of the bars indicated mean values. **(B)** SCGE assays highlighted significant differences in chromatin condensation between LCLs carrying the p.(Asp551Gly) amino acid change and control cells after 20 minutes of electrophoresis run time (**p < 0.05; two-tailed unpaired Student's t test with Welch's correction), that became more evident after 60 minutes of electrophoresis run time (***p 0.0001; two-tailed unpaired Student's t-test with Welch's correction). DNA migration was quantified as Tail moment values, which is defines as the product between the tail length and the percentage of DNA in the tail. For each point, at least 100 cells were analysed. Values are represented as mean ± SEM of three independent experiments. **(C)** Representative images of nucleoids derived from control LCLs and LCLs from affected subject 1 referred to experiment shown in figure 2B. **(D)** Single and double strand breaks were induced by γ-ray irradiation (2-Gy or 4-Gy). Tail moment values specify the amount of γ-ray-induced DNA damage measured immediately after the treatment. The mutant LCLs showed a higher vulnerability to 2-Gy γ-ray irradiation (***p 0.0006; two-tailed unpaired Student's t-test with Welch's correction). Following 4-Gy treatment, no differences were observed between control and mutant cells, which is likely explained by the observation that the overall damage, especially double strand breaks, prevails on the condensation state of chromatin at high doses of γ-ray irradiation. For each point, at least 100 cells were analysed, and four independent experiments were performed. **(E)** Representative images of nucleoids derived from control LCLs and LCLs from affected subject 1 referred to experiment shown in figure 2D.

Figure 3. TLK2 autism mutations alter the activity and subcellular localization of TLK2.

(A-B) Modelling of the D551G and S617L missense mutations on the crystal structure of the TLK2 PKD. Hydrogen bonds are shown in red dashed lines. **(C)** Representative *in vitro* kinase assays with Strep-purified TLK2-WT, TLK2-KD (kinase dead; D592V) and indicated missense variants. Autophosphorylation and substrate (ASF1A) phosphorylation is shown. Coomassie is shown as loading control for ASF1A. **(D)** Quantification of n=3 independent kinase assays. Individual results (circles) are shown for each assay on purified ASF1A substrate or affinity purified TLK2 autophosphorylation relative to corresponding TLK2-WT and bars depict mean with SEM. **(E)** Representative immunofluorescence microscopy of overexpressed TLK2 in AD-293 cells is shown, indicating the 2 main localization patterns observed. The nuclear localization image corresponds to TLK2-WT, while the perinuclear localization to TLK2-S617L. Scale bar = 20 µM. **(F)** Quantification of TLK2 localization patterns for WT and indicated missense variants. Ten random fields were scored in 2 (D551G) or 3 (WT and S617L) biological replicates. Bars depict

mean with SEM. Statistical significance was determined using an unpaired t test with Welch's correction (****P<0.0001, ***P<0.001, **P<0.01, *P<0.05).

Figure 4. BioID based analysis of the proximal interactome of TLK2.

(A) Western blot of AD-293 lysates expressing BioID constructs: N-FLAG-BirA alone or fused to the indicated TLK2 allele. Detection with anti-TLK2, anti-FLAG or Streptavidin-HRP are shown. Ponceau stained nitrocellulose membrane is shown as a loading control. **(B)** Network clustering of all prey hits with a SAINT score of > or = to 0.7 in TLK2-WT samples. Physical interactions reported in Biogrid (solid lines) and functional interactions (dashed lines) reported in STRING are indicated[188,189]. Functionally enriched clusters are indicated by color coding, Bait, TLK2 substrates, proteins found in the SFARI/DECIPHER gene database (yellow fill) or proteins enriched on nascent DNA/replication forks are indicated (red font). **(C-D)** Scatterplots of average spectral counts (Log_2 transformed) of bait and prey proteins identified with the TLK2-D551G and TLK2-S617L alleles compared to TLK2-WT. Previously identified TLK2 interactors, as well as proteins enriched on replication forks or found in the SFARI/DECIPHER databases are indicated (see legend).

Figure 5. Missense variants alter the proximal interactome of TLK2.

(A) Dotplot of prey proteins with a SAINT score of > or = to 0.7 with any of the 3 baits generated using ProHits-viz[190]. Average spectral counts (SC), relative abundance and SAINT score ranges are indicated, as well as proteins enriched on replication forks [171] or found in the SFARI/DECIPHER databases (see legend). Additional details are provided in table S8.

Figure 6: Validation of proximal interactions with CHD7 and CHD8.

(A) Western blots of the indicated proteins from Strep-AP lysates from AD-293 cells transfected with the indicated BioID construct and supplemented with biotin. Input levels are shown and ponceau stained blots provided as a loading control. Representative data from 2 biological replicates is shown.

Figure 7. *MRD57* is associated with a detectable episignature in patients' DNA

(A-B) Differentially methylated probes were selected. **(A)** p-value distribution and (**B)** volcano plot of the mean methylation differences and p-values are shown, where selected probes are in red. (**C)** Hierarchical clustering of training data shows *TLK2* samples grouping separately from unaffected controls. (**D)** MDS plot also shows clear separation of cases from controls. **(E-F)** A support vector machine model was used to classify samples and calculate probability scores. Samples were trained against controls plus samples from other neurodevelopmental (NDD) syndromes: 75% of controls and NDD samples were used for training (blue), and 25% for testing (grey).

Figure 1

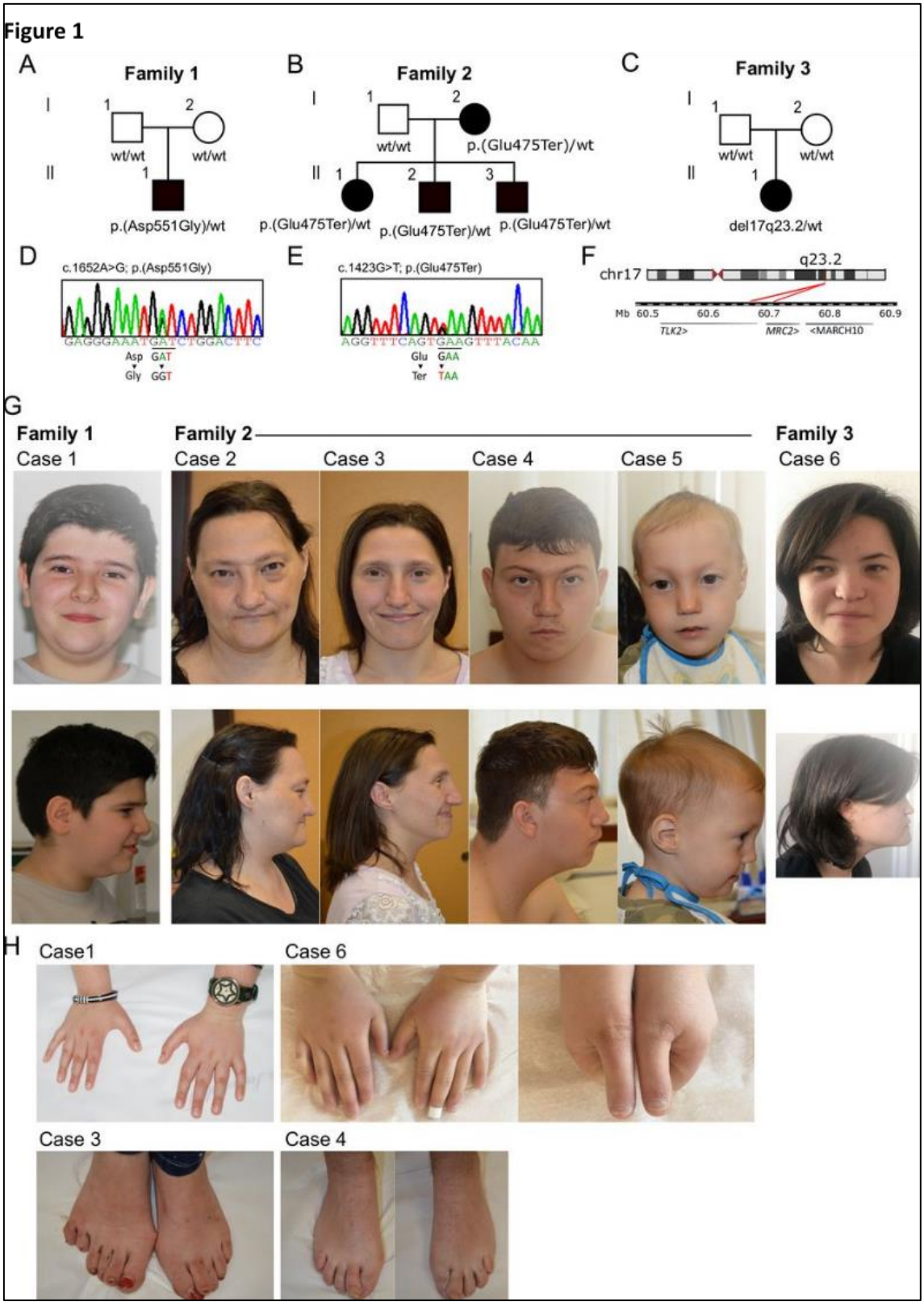

A
Family 1
I
1
2
wt/wt
wt/wt
II
1
p.(Asp551Gly)/wt
B
Family 2
I
1
2
wt/wt
p.(Glu475Ter)/wt
II
1
2
3
p.(Glu475Ter)/wt
p.(Glu475Ter)/wt
p.(Glu475Ter)/wt
C
Family 3
I
1
2
wt/wt
wt/wt
II
1
del17q23.2/wt
D
c.1652A>G; p.(Asp551Gly)
GAGGGAAATGATCTGGACTTC
Asp GAT
Gly GGT
E
c.1423G>T; p.(Glu475Ter)
AGGTTTCAGTGAAGTTTACAA
Glu GAA
Ter TAA
F
chr17
q23.2
Mb 60.5 60.6 60.7 60.8 60.9
TLK2>
MRC2>
<MARCH10
G
Family 1
Case 1
Family 2
Case 2
Case 3
Case 4
Case 5
Family 3
Case 6
H
Case1
Case 6
Case 3
Case 4

Figure 2

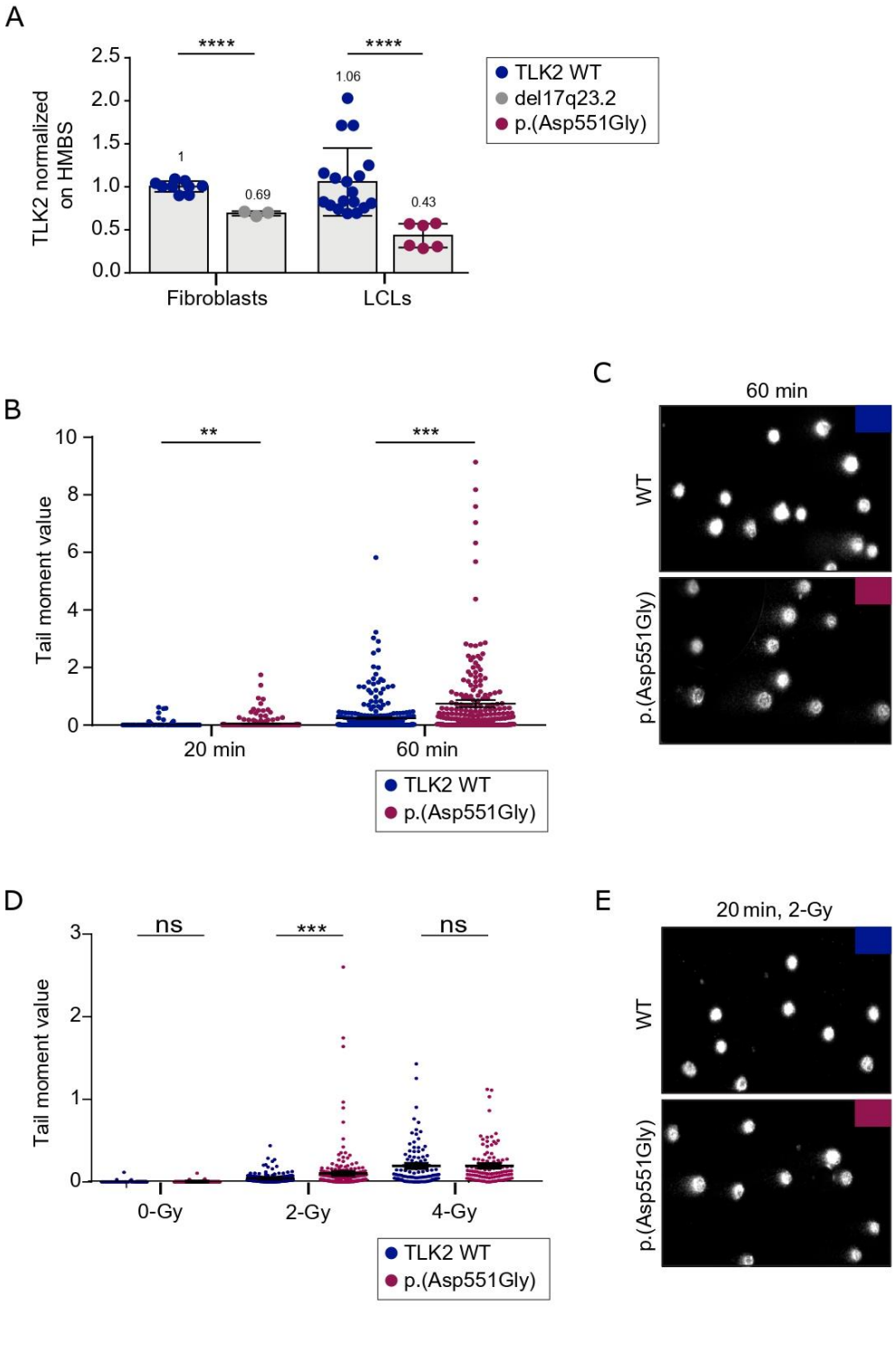

A
****
****
2.5
2.0
1.5
1.0
0.5
0.0
TLK2 normalized on HMBS
1
0.69
1.06
0.43
Fibroblasts
LCLs
TLK2 WT
del17q23.2
p.(Asp551Gly)
B
**
***
10
8
6
4
2
0
Tail moment value
20 min
60 min
TLK2 WT
p.(Asp551Gly)
C
60 min
WT
p.(Asp551Gly)
D
ns
***
ns
3
2
1
0
Tail moment value
0-Gy
2-Gy
4-Gy
TLK2 WT
p.(Asp551Gly)
E
20 min, 2-Gy
WT
p.(Asp551Gly)

Figure 3

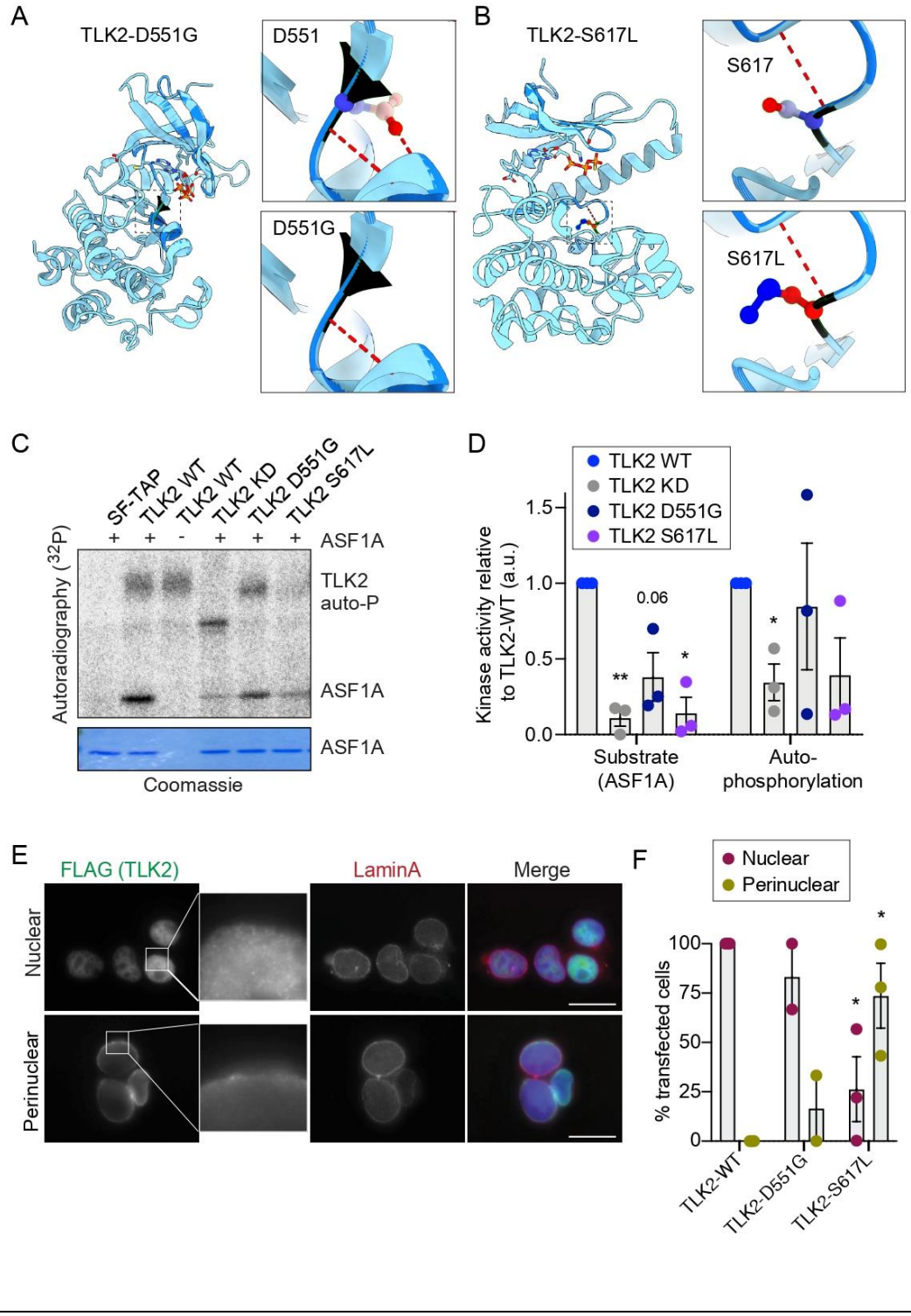

A
TLK2-D551G
D551
D551G
B
TLK2-S617L
S617
S617L
C
SF-TAP
TLK2 WT
TLK2 WT
TLK2 KD
TLK2 D551G
TLK2 S617L
+ + - + + +
ASF1A
Autoradiography (32P)
TLK2 auto-P
ASF1A
ASF1A
Coomassie
D
TLK2 WT
TLK2 KD
TLK2 D551G
TLK2 S617L
Kinase activity relative to TLK2-WT (a.u.)
1.5
1.0
0.5
0.0
0.06
**
*
*
Substrate (ASF1A)
Auto-phosphorylation
E
FLAG (TLK2)
LaminA
Merge
Nuclear
Perinuclear
F
Nuclear
Perinuclear
% transfected cells
100
75
50
25
0
*
*
TLK2-WT
TLK2-D551G
TLK2-S617L

Figure 4

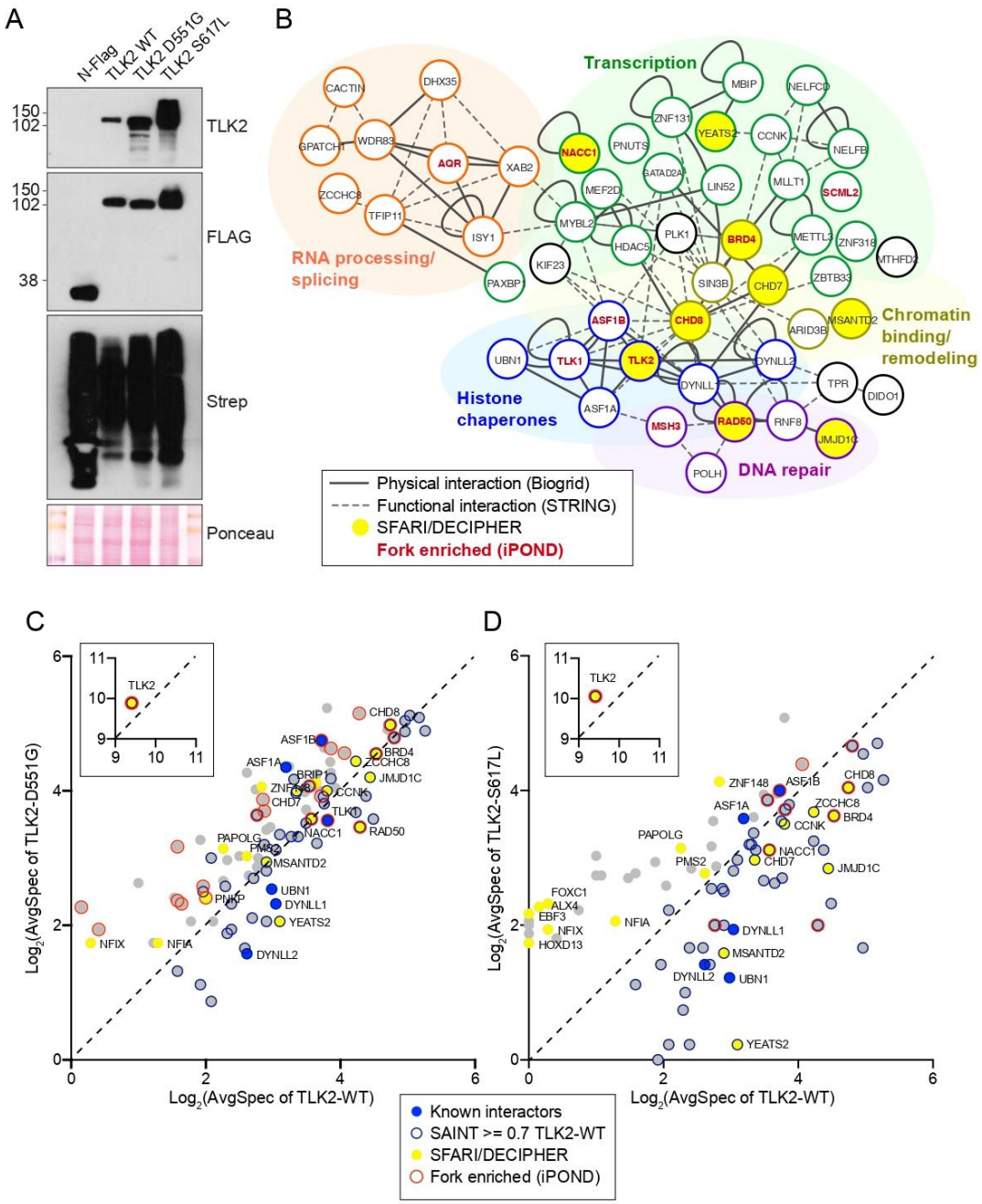

A
N-Flag
TLK2 WT
TLK2 D551G
TLK2 S617L
150
102
TLK2
FLAG
38
Strep
Ponceau
B
Transcription
RNA processing/ splicing
Chromatin binding/ remodeling
Histone chaperones
DNA repair
Physical interaction (Biogrid)
Functional interaction (STRING)
SFARI/DECIPHER
Fork enriched (iPOND)
C
D
Log2(AvgSpec of TLK2-D551G)
Log2(AvgSpec of TLK2-S617L)
Log2(AvgSpec of TLK2-WT)
Known interactors
SAINT >= 0.7 TLK2-WT
SFARI/DECIPHER
Fork enriched (iPOND)

Figure 5

A

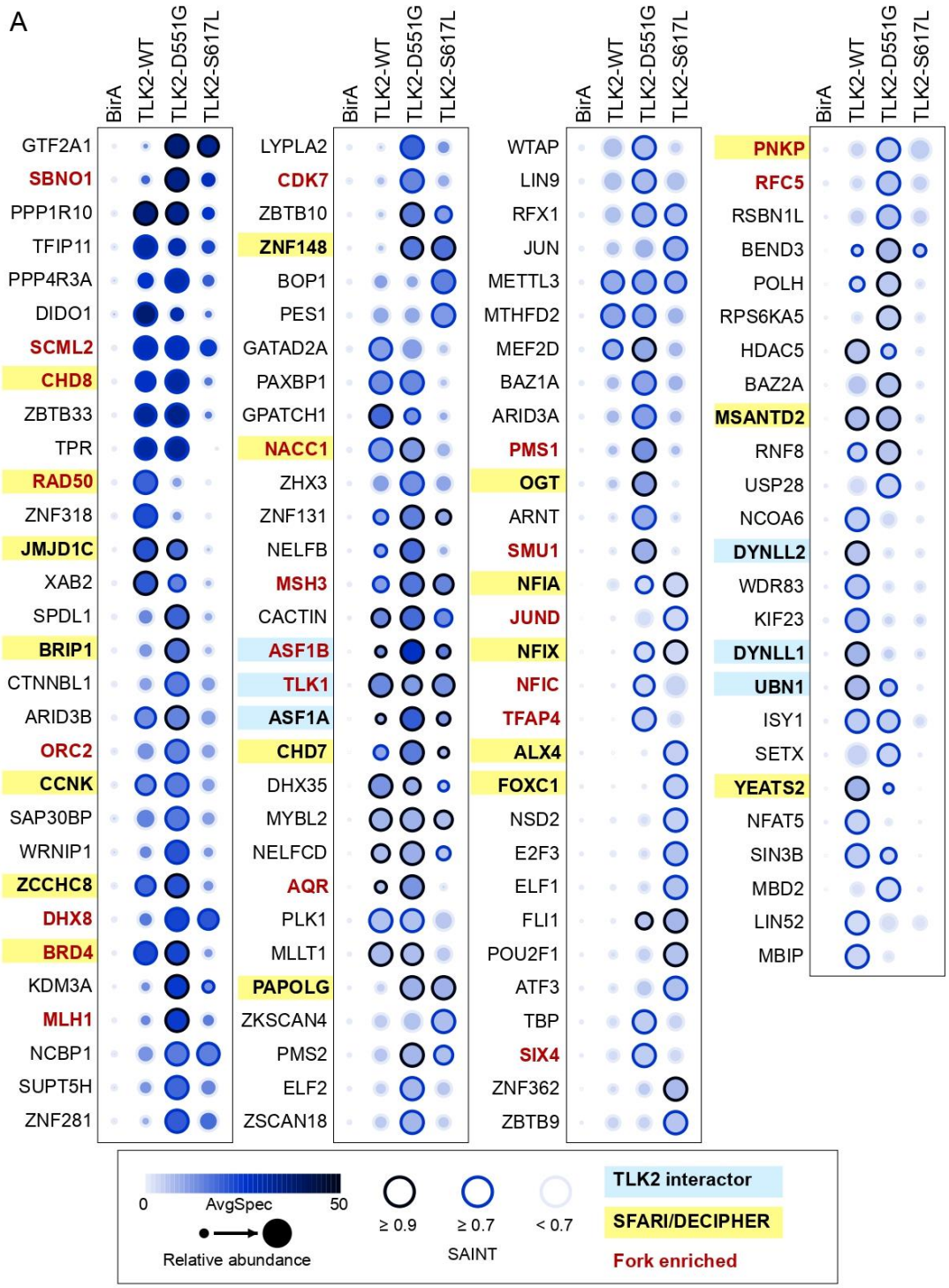

BirA
TLK2-WT
TLK2-D551G
TLK2-S617L
GTF2A1
SBNO1
PPP1R10
TFIP11
PPP4R3A
DIDO1
SCML2
CHD8
ZBTB33
TPR
RAD50
ZNF318
JMJD1C
XAB2
SPDL1
BRIP1
CTNNBL1
ARID3B
ORC2
CCNK
SAP30BP
WRNIP1
ZCCHC8
DHX8
BRD4
KDM3A
MLH1
NCBP1
SUPT5H
ZNF281
LYPLA2
CDK7
ZBTB10
ZNF148
BOP1
PES1
GATAD2A
PAXBP1
GPATCH1
NACC1
ZHX3
ZNF131
NELFB
MSH3
CACTIN
ASF1B
TLK1
ASF1A
CHD7
DHX35
MYBL2
NELFCD
AQR
PLK1
MLLT1
PAPOLG
ZKSCAN4
PMS2
ELF2
ZSCAN18
WTAP
LIN9
RFX1
JUN
METTL3
MTHFD2
MEF2D
BAZ1A
ARID3A
PMS1
OGT
ARNT
SMU1
NFIA
JUND
NFIX
NFIC
TFAP4
ALX4
FOXC1
NSD2
E2F3
ELF1
FLI1
POU2F1
ATF3
TBP
SIX4
ZNF362
ZBTB9
PNKP
RFC5
RSBN1L
BEND3
POLH
RPS6KA5
HDAC5
BAZ2A
MSANTD2
RNF8
USP28
NCOA6
DYNLL2
WDR83
KIF23
DYNLL1
UBN1
ISY1
SETX
YEATS2
NFAT5
SIN3B
MBD2
LIN52
MBIP
0
AvgSpec
50
Relative abundance
≥ 0.9
≥ 0.7
< 0.7
SAINT
TLK2 interactor
SFARI/DECIPHER
Fork enriched

Figure 6

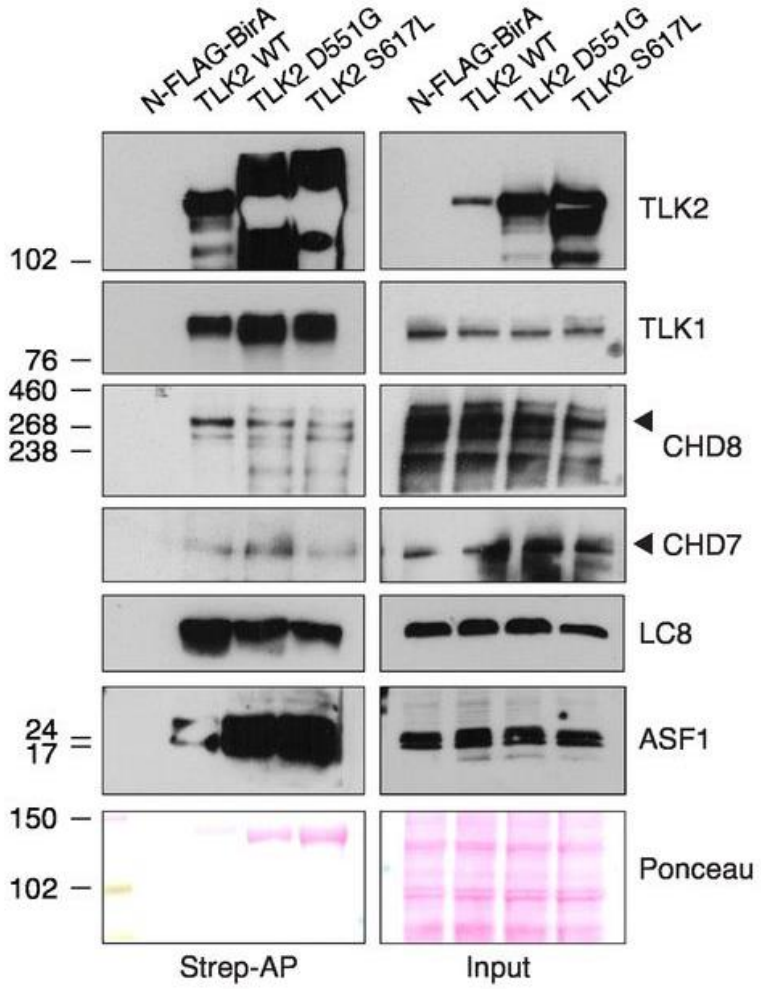

N-FLAG-BirA
TLK2 WT
TLK2 D551G
TLK2 S617L
N-FLAG-BirA
TLK2 WT
TLK2 D551G
TLK2 S617L
TLK2
TLK1
CHD8
CHD7
LC8
ASF1
Ponceau
102
76
460
268
238
24
17
150
102
Strep-AP
Input

Figure 7

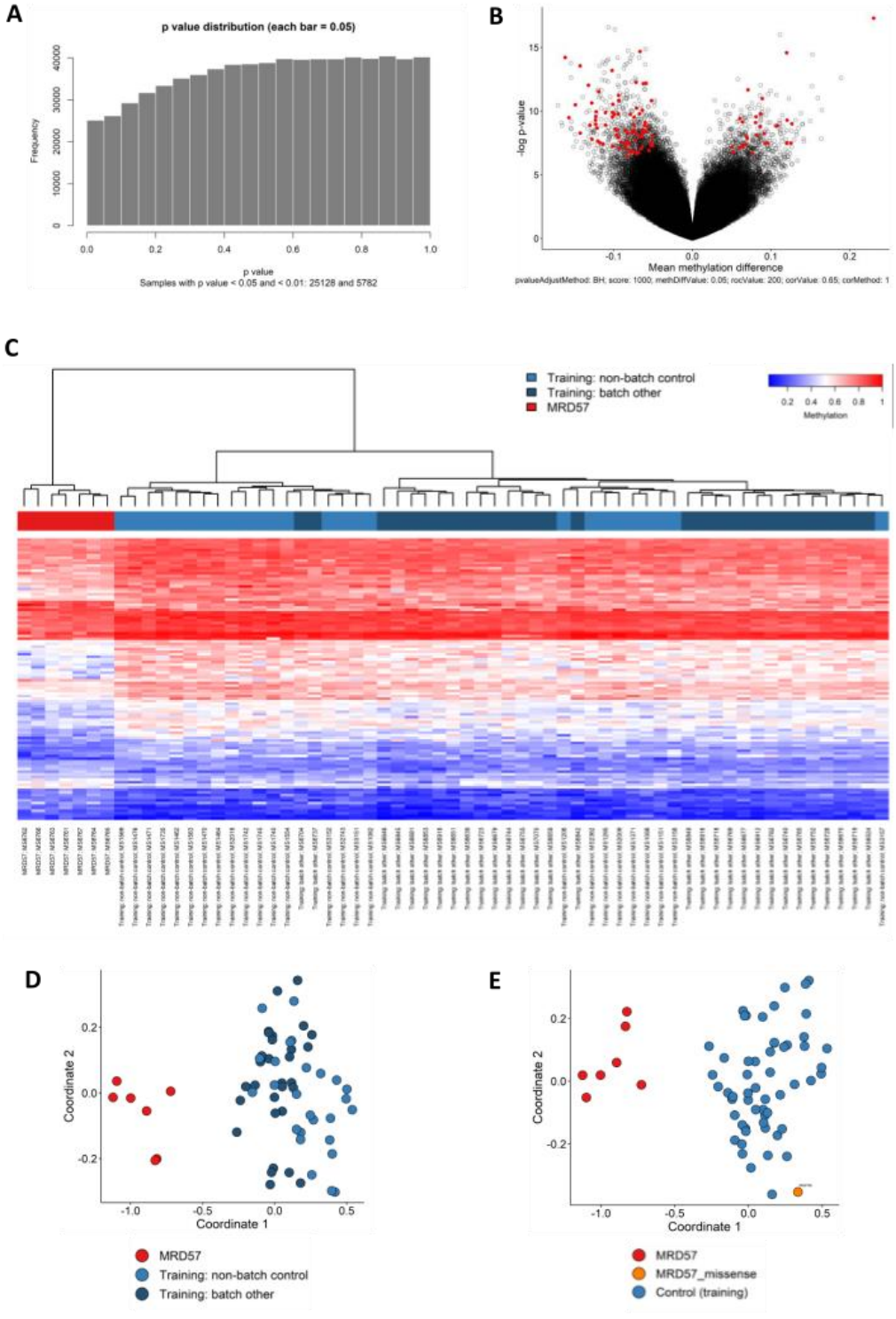
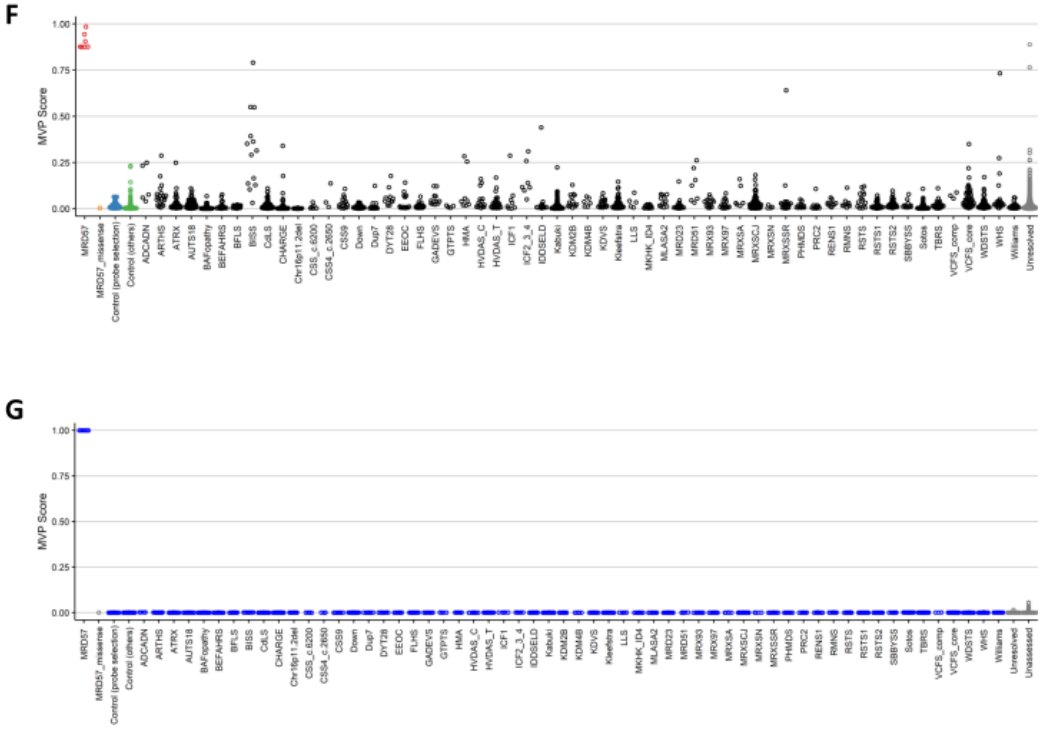

A
p value distribution (each bar = 0.05)
Frequency
p value
Samples with p value < 0.05 and < 0.01: 25128 and 5782
B
-log p-value
Mean methylation difference
C
Training: non-batch control
Training: batch other
MRD57
Methylation
D
Coordinate 1
Coordinate 2
MRD57
Training: non-batch control
Training: batch other
E
Coordinate 1
Coordinate 2
MRD57
MRD57_missense
Control (training)

Figure 7 - continued



Chapter 4: Novel phenotypes associated to known genes

Within the NeuroWES project, we found different cases of novel phenotypes associated to known genes. An example is *WDR26* gene for which, at the time of our analysis, only 18 patients were described, and no information on the natural history of the disease was reported.

We found two additional cases with variants in *WDR26* gene and carefully reported the clinical characterization expanding the knowledge on the phenotype associated to this gene (4.1).

In other cases, partially overlapping but distinct phenotypes were identified in our patients. An example is the identification of a patient with two variants in *PI4KA* (NM_058004.4) gene [c.1414A>C; p.(Ser472Arg) paternal and c.355C>T; p.(Arg119Trp) maternal]. Our patient showed severe global developmental delay, severe intellectual disability, delayed gross motor development, unsteady gait, delayed speech and language development, impaired social interactions, abnormal myelination, stereotypy, and abnormal muscle tone. At the time of our analysis, *PI4KA* gene was only associated with "Polymicrogyria, perisylvian, with cerebellar hypoplasia and arthrogryposis" (OMIM #616531), and mutations were reported only in three fetuses from the same family. Thanks to GeneMatcher, we found other patients with a phenotype similar to our case, leading to the conclusion that variants in this gene can cause two separate disorders with some overlapping features and different levels of severity[37]. Another example, described below, is the identification of a family with variants in *KCNK18* gene which, at the time of the analysis, was only associated with a form of migraine (OMIM #613656), but that we found mutated in three affected patients with a neurodevelopmental disorder[191](4.2).

Phenotype heterogeneity is also perfectly showed by *SUPT5H*: this gene has been previously associated with a β-thalassemia trait in four families[192]. We found a *SUPT5H* variant in a patient with NDD and increased HbA_2; GeneMatcher query allowed us to find eleven additional patients with *SUPT5H* variants associated with NDD. This ongoing study is better described at the end of this chapter (4.3).

4.1. WDR26 variants: expanding the clinical phenotype of Skraban-Deardorff syndrome

Haploinsufficiency of the WD Repeat Domain 26 (*WDR26*) gene has been associated with Skraban-Deardorff syndrome (MIM #617616), an ultrarare autosomal dominant disorder characterized by intellectual disability (ID), developmental delay (DD), seizures and common facial features, including depressed nasal root and broad nasal tip, anteversion of nares, prominent smile, widely spaced teeth, and gingival abnormalities[193,194]. Happy and friendly demeanour and wide-based and/or stiff-legged gait are often observed[195].

WDR26 is an evolutionary conserved gene and is highly intolerant to loss–of-function variation (pLI = 1.0; Z-score = 3.58, gnomAD ver.2.1.1). This suggests haploinsufficiency as underlying mechanism. The exact pathogenic mechanisms are still unclear, although a role in MAPK, Wnt, and PI3K signalling pathways has been suggested [196,197].

To date, only 15 affected patients (ten females and five males) have been identified and clinically described [195]. Another case with a nonsense variant [198], and two cases with missense variants have been identified by trio-ES, but without detailed descriptions of the clinical phenotypes [157,199]. Notably, 12 of the 18 *WDR26* causative variants previously reported were truncating, further supporting haploinsufficiency as the molecular mechanism of the disease (Figure 1A).

Here, we report two novel loss-of-function *de novo* variants in *WDR26* in one paediatric and one adult patient with Skraban-Deardorff syndrome. Detailed clinical characterization of the two cases allowed us to identify new features of Skraban-Deardorff syndrome. Furthermore, we profile the history of this disease and provide insights useful for the management of affected individuals.

Methods

Whole exome sequencing, prioritization, and variant calling

DNA was extracted from total blood using the ReliaPrep Blood gDNA Miniprep kit (Promega, Madi-son, WY, USA) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientifics, Waltham, MA, USA).

Array-CGH was performed using a 60K whole-genome oligonucleotide microarray (Agilent Technologies, Santa Clara, California, USA).

Patients were enrolled in the Autism Sequencing Consortium (ASC) project and their gDNA samples were sequenced at the Broad Institute on Illumina HiSeq sequencers as previously described [6,26].

ES raw data of the trio were processed and analyzed using an in-house implemented pipeline previously described [27,28] which is based on the GATK Best Practices [29]. The UCSC GRCh37/hg19 version of genome assembly was used as a reference for reads alignment by means of BWA-MEM [30] tool and the subsequent variant calling with HaplotypeCaller (GATK v3.7)[29]. We used SnpEff v.4.3 [31] and dbNSFP v.3.5 [32] tools for variants functional annotation, including Combined Annotation Dependent Depletion (CADD) v.1.3 [33], Mendelian Clinically Applicable Pathogenicity (M-CAP) v.1.0 [200] and Intervar v.0.1.6 for functional impact prediction [34]. Thereby, the analysis was narrowed to variants which affect coding sequences or splice site regions. Moreover, high-quality variants were filtered against public databases (dbSNP150 and gnomAD ver.2.0.1) so that

only variants with unknown frequency or having MAF <0.1%, as well as variants occurring with frequency <1% in our population-matched database (~2000 exomes) were considered.

A further variant stratification in conformity with the American College of Medical Genetics and Genomics (ACMG) guideline [34], also considering mode of inheritance and functional *in-silico* prediction of impact, allowed us considering the final set of variants for possible associations with the phenotype. All variants are referred to GRCh37 annotation and to NM_025160.6.

Identified variants were confirmed by Sanger sequencing using standard conditions and the following primers: 5'- ctcctccgtggtggtagtgg; 5'- gacgactccccgttctgg (for c.69dupC, p.(Gly24ArgfsTer48)) and 5'-gagtctcttgaccccaggcg; 5'-cagtttcctcatctgactgcagg (for c.1076G>A, p.(Trp359Ter)).

In silico prediction of variants impact

Variants were analysed with the VarSome tool [159] as a starting point for further analysis. This allowed evaluation of at least nine *in silico* predictors simultaneously. Variants frequencies were evaluated using Genome Aggregation Database (GnomAD) Browser version 2.1.1.

Ethical committee

Informed consent was obtained from participating families and the study protocol was approved by the internal Ethics Committee of University of Turin (n. 0060884) and University of Skopje (n. 03-6116/7), according to the Declaration of Helsinki.

Results

Clinical description of patient 1

Patient 1 was an 8-year-old Caucasian female, born at 40th weeks of gestation via spontaneous vaginal delivery as a second child from a non-consanguineous parents from Macedonia (Apgar 9/10; birth weight 3,000 g, 50th centile; birth length 51 cm, 50th centile) (Supplementary Figure 1A). The pregnancy was uncomplicated and prenatal ultrasounds were normal; the family history was unremarkable.

The patient could hold her head at 5 months, sat at 7 months and walked at 17 months; the control of the sphincters was reached at 4 years of age. Language delay was also reported. Previous evaluation included normal basic metabolic screenings and normal karyotype.

Her medical history was significant for bilateral foot deformity with *metatarsus varus* (Figure 1B), recurrent serous otitis media requiring tympanostomy tubes, hypertrophic adenoids, and food allergies at age of 7 years old. Her hearing Brain Stem Evoked Response Audiometry (BERA) examination showed tympanometry curve type "B" on the right, and type "C" on the left ear.

Beside the report of neonatal hypotonia, the child had normal muscle strength accompanied by dyspraxia and wide-based gait.

At 4 years old, she was evaluated for possible absence seizures (described by the mother as staring episodes with frequent blinking and mouth chewing) however the sleep-deprived EEG was inconclusive. A second EEG assessment (with hyperventilation/photic stimulation) could not be performed to obtain a definitive diagnosis [201].

The patient had a playful and affectionate demeanour yet became shy with strangers. Aggressiveness was occasionally observed. At the last evaluation, she had not acquired significant verbal language, her communication was limited to a few words, and she often communicated with gestures. The presence of an adult, especially her mother, was required for every activity. Memory could not be evaluated because of lack of verbalization and expressive aphasia.

The proband had prominent brows with long lashes, large irises and upslanted palpebral fissures, depressed nasal root and broad nasal tip, mild anteverted nares, full cheeks, slightly long philtrum, wide mouth, widely spaced teeth, abnormal gums, and gingival hyperplasia (Figure 1B). Pictures from childhood along with computational comparison with previously reported cases are provided in Figures 1B and D.

Clinical description of patient 2

Patient 2 was enrolled in the ASC project at 25 years old, after a long diagnostic odyssey, and received her diagnosis 4 years later. A detailed description of the medical reports performed tests and administered drugs is provided in Figure 2 and Table S1.

She was the second child (Supplementary Figure 1A) of non-consanguineous Caucasian parents from Italy, born at 38th gestational weeks after a complicated pregnancy due to placental insufficiency and sclerosis of the chorionic villi. Parents were healthy without history of neuropsychiatric disorders, beside report of ID of the maternal first cousin.

At postnatal day 13, a slight increase in the volume of the heart left cavity, the right ventricle and of the pulmonary artery were reported together with a perimembranous ventricular septal defect. Since then, the patient was clinically followed due to severe developmental and motor delays.

At two years old, computed tomography (CT) scan (Figure 2B) and magnetic resonance imaging (MRI, not available) exam showed reduced dimension of the splenium of the corpus callosum and periventricular leukomalacia, with a final diagnosis of suspected connatal malformative encephalopathy. Routine metabolic screening was normal.

The first neuropsychiatric evaluation was performed at 3 years of age, when several manual and oral motor disabilities (semi-open mouth, frequent movement of protrusion of the tongue without drooling, absent chewing, oral-motor apraxia) were reported. She also presented two types of generalized epilepsy: motor tonic-clonic, and non-motor absence crisis (when seizures were more frequent, the patient fell asleep or remained confused at awake) with a duration of about 10 seconds each.

Walking began at 3 years of age, with slow and unstable wide-based stiff-legged gait, often requiring support from a wheelchair. Fine motor skills were particularly compromised, as well as her muscle coordination presented with dysmetria and dyspraxia. Dysphagia was identified during the infancy and continued lifelong, with the preferential consumption of liquid foods.

At 22 years old, her brain MRI showed a modest enlargement of cerebral ventricles with alterations of the periventricular white matter suggestive of neonatal hypoxia. The corpus callosum was modestly thinner, while pineal gland was normal (Figure 2B).

As an adult, she showed good spatial orientation only in familiar places and became agitated in new ones. She was unable to discriminate right from left. Interestingly she showed hypersensitivity to smell and touch, and she was not able to filter different sensorial stimuli.

Throughout her life the patient had a poorly organized recreational activity, extremely poor social interaction, and deep anxiety. The Griffith scale [202] reported medium severity global delay, reduced ability of visual attention, impaired visual monitoring, and ability of considering many things at the same time. The attention span was limited, due to extremely low motivation. Memory could not be evaluated because of lack of verbalization. Taking in consideration all these features a diagnosis of autism spectrum disorder (ASD) associated with language disorder and severe intellectual disability was reached.

Concerning patient's perspective with health outcome, a complete habilitative daily program, based on behavioral therapy in a small ASD group, has been provided from Adult autism center and Social Agency to improve social skill but very little improvement has been obtained, so the treatment was further based on care and support [203].

She had a happy and friendly demeanour for the most part, but could become agitated and scared, especially in case of new situations.

She presented good facial mimic with lively and expressive look, wide mouth with thin and elongated lips, decreased Cupid's bow, abnormal gingiva, widely spaced teeth, as well as large and protruding tongue. Skeletal findings included bilateral clinodactyly of fifth finger and partial agenesis of distal part of the second finger of left hand. Pictures from childhood to adulthood and a computational comparison with already reported cases are provided in Figure 1B and C.

A detailed clinical characterization of the two patients together with the clinical phenotype of the previously reported Skraban-Deardorff patients is presented in Figure 3 and divided by different systems.

Molecular findings

Our probands were part of a large cohort of 686 predominantly Caucasian patients affected by ASD with or without ID collected for genetic screening and analysed by whole exome sequencing as previously reported [6,26].

In patient 1, we identified the *de novo* frameshift variant c.69dupC, p.(Gly24ArgfsTer48) in *WDR26*. The single nucleotide insertion occurred in the first exon of 14 coding exons, causing a frameshift variant that is predicted to trigger non-sense mediated decay. In patient 2, we found the *de novo* nonsense variant c.1076G>A, p.(Trp359Ter) in *WDR26* introducing a premature stop codon, that predicted to activate non-sense mediated decay. Both variants were classified as "pathogenic" according to ACMG criteria (PVS1, PS2, PM2) [34]. Moreover, concerning the *de novo* nonsense variant identified in patient 2, eight out of nine *in silico* tools predicted this change as damaging, adding the PP3 supporting criteria to ACMG classification. Both variants were absent in the Genome Aggregation Database (gnomAD ver 2.1.1) or in the 1000 Genomes browser. No further clinically relevant variants were identified.

Discussion

We report one paediatric and one adult case with *de novo* loss-of-function variants in *WDR26* associated with the ultra-rare Skraban-Deardorff syndrome. Our patients showed overlapping phenotype with reported cases. However, patient 2 had several additional features, further expanding the Skraban-Deardorff syndrome-related phenotype (Table S2). It is interesting to note that many of these newly observed characteristics appeared after puberty, suggesting a possible evolution of the disease that has not been reported yet. Indeed, almost all of the described cases are paediatric, with the exception of two cases of 21 and 34-years old [195].

The consensus phenotype for Skraban-Deardorff syndrome includes intellectual disability, delayed or absent speech, autistic features, dysphagia, wide-base and/or stiff-legged gait, skeletal anomalies, recurrent otitis media and brain MRI anomalies (prevalently affecting corpus callosum). The similarity of facial phenotype is striking, underlying the consistent facial gestalt in this disorder and the relevance of dysmorphological evaluation in the diagnostic path of ASD patients.

Notably, seizures have been observed in all cases, mainly generalised motor, and non-motor absences seizures. The latter (also called *petit mal seizures*) are hard to diagnose, especially in non-verbal children [204]. Other condition that might further bias the diagnosis are the frequent otitis media that could give the impression that the child is not responsive due to hearing problems. For this reason, we suggest that a more accurate phenotyping of the type of epilepsy would provide relevant insights for a more effective management of these patients.

Beside the absence epilepsy in patient 2, a frequent sleep disturbance and tactile and olfactory hypersensitivity were reported. It is intriguing to observe that these clinical features can be tracked to the thalamus [205,206]. Indeed, thalamus transmits sensory information with a rhythmic pattern in the thalamocortical tract, typically tuned down during sleep. It appears that in absence epilepsy and in people having sleeping difficulties, this thalamocortical pattern is changed and amplified into a distortion range [207]. This information could help in further understanding the mechanisms underlying the disorder and open new avenues for treatment [208,209].

Hypotonia has been reported in 9 out of 12 cases [195]. Surprisingly, our patients showed remarkable muscle strength and often hypertonia; hypertonic limbs with brisk reflexes were also previously reported for one case [195]. We can speculate that this ambiguous muscle tone phenotype, as well as the diverse type of walking, could be related to the brain MRI findings, predominantly affecting the white matter such as corpus callosum anomalies [210,211], enlarged ventricles, white matter volume loss, and mild cerebellar hypoplasia [195]. Furthermore, the combination of wide-based gait and spastic and/or stiff-leg gait are often reported. They indicate the involvement of the corticospinal tract and the upper motor neurons. The inclusion of a spinal MRI in Skraban-Deardorff patients may elucidate the origin of these symptoms.

Skeletal findings included metatarsus varus in patient 1, and bilateral clinodactyly of the fifth finger, and partial agenesis of distal part of the second finger of left hand in patient 2. Moreover, as described for one patient [195], a foot malformation was present (right flat and supine foot and left cavus and metatarsus varus) (Figure 1B). Steopathia striata of the distal femurs, *pes cavus* or hip dysplasia that were described in some cases, where not identified in our patients.

The immune system seems also often affected in Skraban-Deardorff with a variety of presentations starting from hypersensitivity reactions to food and airborne allergies, to weak immune response presented by frequent *otitis media*.

Patient 2 showed also additional features such as hypothyroidism, renal acidosis, and various gynaecological abnormalities. Given the facial and behavioural similarity between individuals with Skraban-Deardorff and individuals diagnosed with Angelman syndrome (MIM #105830),

Kabuki syndrome (MIM #147920; reports of hypothyroidism; premature thelarche) or Cornelia de Lange syndrome (MIM #122470; renal anomalies), it is reasonable that Skraban-Deardorff phenotype might result from altered chromatin regulation, as suggested by [195].

WDR26 is located on chromosomal region 1q42, which can be disrupted in 1q41q42 microdeletion syndrome, a chromosomal anomaly resulting in severe developmental delay and/or intellectual disability, typical facial dysmorphic features, brain anomalies, seizures, cleft palate, clubfeet, nail hypoplasia and congenital heart disease [212]. Many features of this syndrome overlapped with the one observed for *WDR26*-related disease and a comparison has already been provided previously [195].

WDR26 gene encodes a WD40 repeat protein 26 (Wdr26), a member of the WD repeat (WDR) protein family. The WDR is one of the most abundant protein-protein interactions domain, with its central peptide-binding pocket [213].

The mechanism by which *WDR26* haploinsufficiency leads to human developmental disorders is unknown. However, multiple roles have been proposed for *WDR26* and some studies suggest that it could play several roles in the regulation of MAPK, Wnt, and PI3K pathway signalling.

WDR26 is important for the degradation of β-catenin in the canonical Wnt pathway, and it has been suggested that the gene is important for the Wnt signalling during embryogenesis [214]. Alterations in Wnt have been implicated in both neurodevelopmental and neurodegenerative disorders. Interestingly, the impairment of this pathway has also been involved in cardiovascular, kidney, lung, allergy, skeletal and oral diseases [196]. WDR26 is also involved in the regulation of Gβγ-mediated PI3K/AKT signalling, acting as a scaffolding protein to promote the interactions between Gβγ, PI3Kβ and AKT2[197]. The PI3K-Akt-mTOR cascade has been broadly implicated in both neurodevelopmental and neuropsychiatric diseases[215]. Therefore, since the *WDR26* haploinsufficiency affects many systems, an intriguing possibility is that the impaired activation of the Wnt and/or PI3K/AKT cascade could be involved in the onset of the disease.

In summary, we report an extended clinical description of two novel Skraban-Deardorff syndrome cases, adding new information about the clinical spectra and the evolvement of the disease from childhood to adulthood and providing a review of the available literature related to the disease. This deep clinical characterization and description of pharmacological treatments might inform management of Skraban-Deardorff patients and give suggestion for further functional characterization of the gene and related disorder.

Figures

Figure 1. WDR26 pathogenic variants and facial features in Skraban-Deardorff syndrome patients.

(A) WDR26 protein structure reported by the National Center for Biotechnology Information (NCBI) (NP_079436.4) contains two domain types: a C-terminal LIS homology domain (CTLH, amino acid residues 156-231; light blue), and several WD repeat domains (yellow); here, six WD domains are shown, but from five to 14 have been estimated. Currently, twenty pathogenic (red)/likely pathogenic (orange) variants (ACMG criteria), including the two here described (p.(Gly24ArgfsTer48) and p.(Trp359Ter) in bold) are reported (five frameshifts, five premature stop codons, seven missense and one splice site variants). **(B)** Our probands' pictures at different ages from childhood to adulthood show the evolution of facial appearance and dysmorphisms. Below feet malformations include flatfoot, forefoot varus and metatarsus varus. **(C)** Facial features analysis by Face2Gene software (FDNA Inc., Boston, USA) [111]. Heat map (on the left) were automatically generated by Face2Gene and shows overlapping features with 15 published Skraban-Deardorff syndrome cases [195] and compared with our Patient 1 at 7 yrs. and Patient 2 at 25 yrs.; overlapping facial regions are indicated by the coloured halo from red to blue. A computational face comparison with already reported cases carrying *WDR26* loss-of-function variants is shown on the right.

Figure 2. Timeline of the Patient 2 diagnostic odyssey, treatments and neuroradiology.

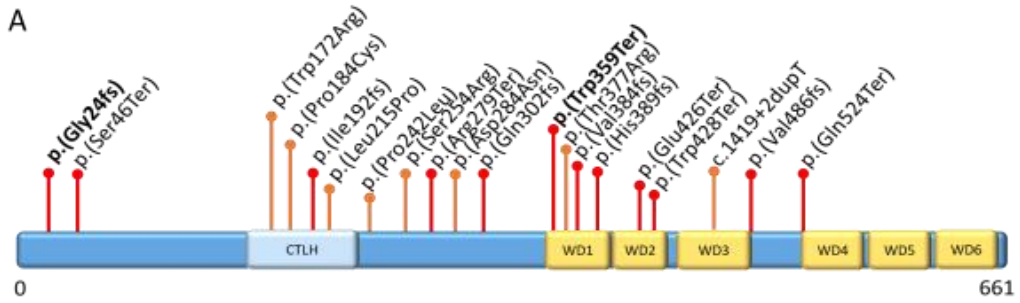
(A) Summary of the key medical reports performed tests and administered drugs is reported. The first medical record dates to 13 post-natal days, when cardiovascular defects were noted. The patient was subjected to a long diagnostic odyssey before finding a molecular diagnosis thanks to trio whole exome sequencing, which happened at 29 years old. **(B)** On the left, brain computer tomography (CT) scan (T1 sagittal section) at 2 years of age showed reduced dimension of the splenium of the corpus callosum and periventricular leukomalacia. On the right: At 22 years old, MRI axial T2 flair showed an enlargement of cerebral ventriculi with features resembling neonatal hypoxia (left, white arrows); a modest reduction in corpus callosum size was also apparent (right, white arrow).

Figure 3. Phenotypic comparison between previously reported Skraban-Deardorff syndrome cases and our patients.

Graphical representation of the clinical features previously reported in other Skraban-Deardorff syndrome cases (light blue) and in our patients (pink and magenta). Numbers indicate the patients showing the phenotype, considering a total number of 17 cases (15 reported in Skraban

et al. 2017 work and our cases). Patients reported by [157,198,199] were excluded from this count, as detailed clinical information were not available.

Figure 1


A
P.(Gly24fs)
p.(Ser46Ter)
p.(Trp172Arg)
p.(Pro184Cys)
p.(Ile192fs)
p.(Leu215Pro)
p.(Pro242Leu)
p.(Ser254Arg)
p.(Arg279Ter)
p.(Asp284Asn)
p.(Gln302fs)
p.(Trp359Ter)
p.(Thr377Arg)
p.(Val384fs)
p.(His389fs)
p.(Glu426Ter)
p.(Trp428Ter)
c.1419+2dupT
p.(Val486fs)
p.(Gln524Ter)
CTLH
WD1
WD2
WD3
WD4
WD5
WD6
0
661


B
Patient 1
Patient 2
Skeletal features


C
Our case
Reported cases
Patient 1
Patient 2
Facial feature of Skraban-Deardorff patients

Figure 2

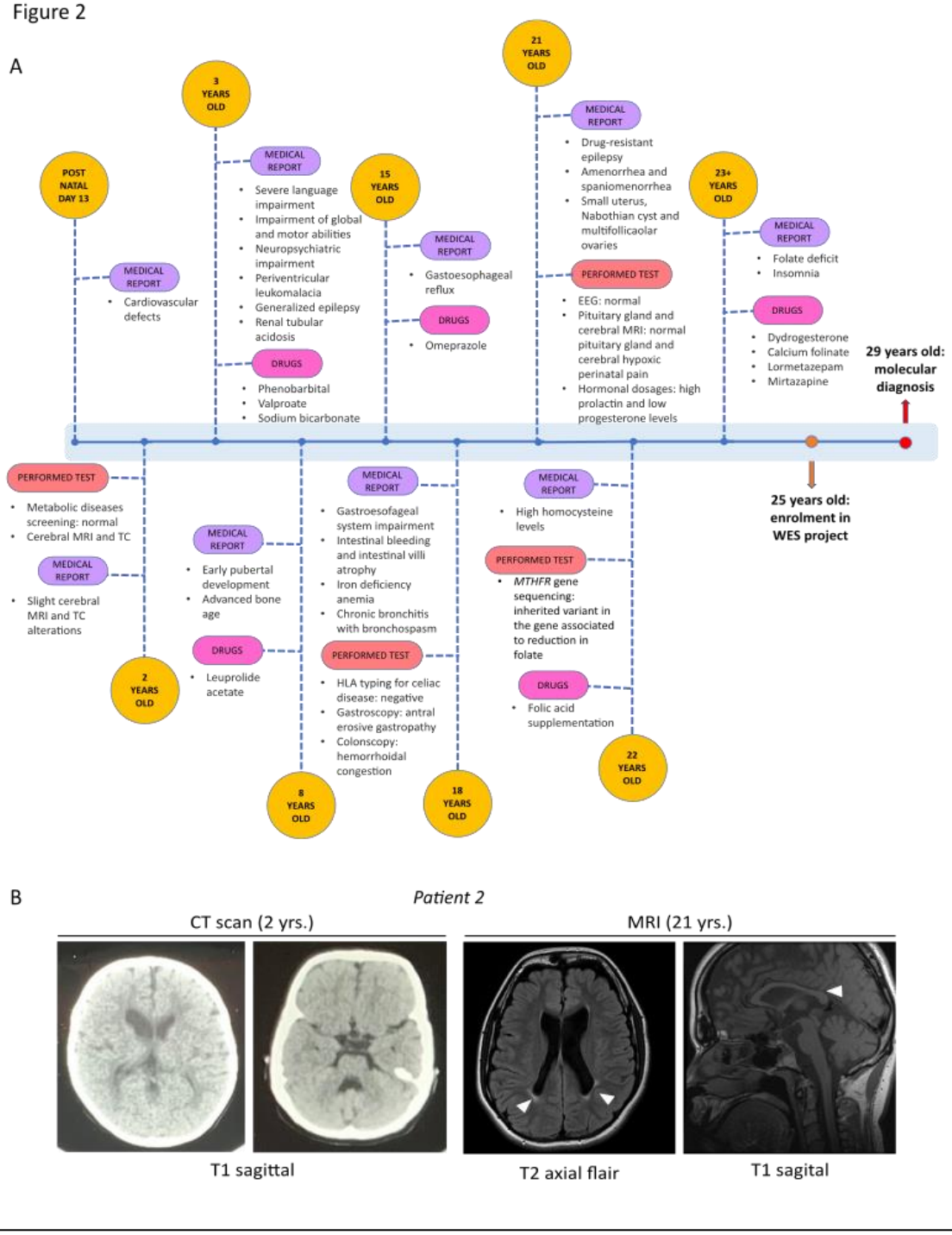
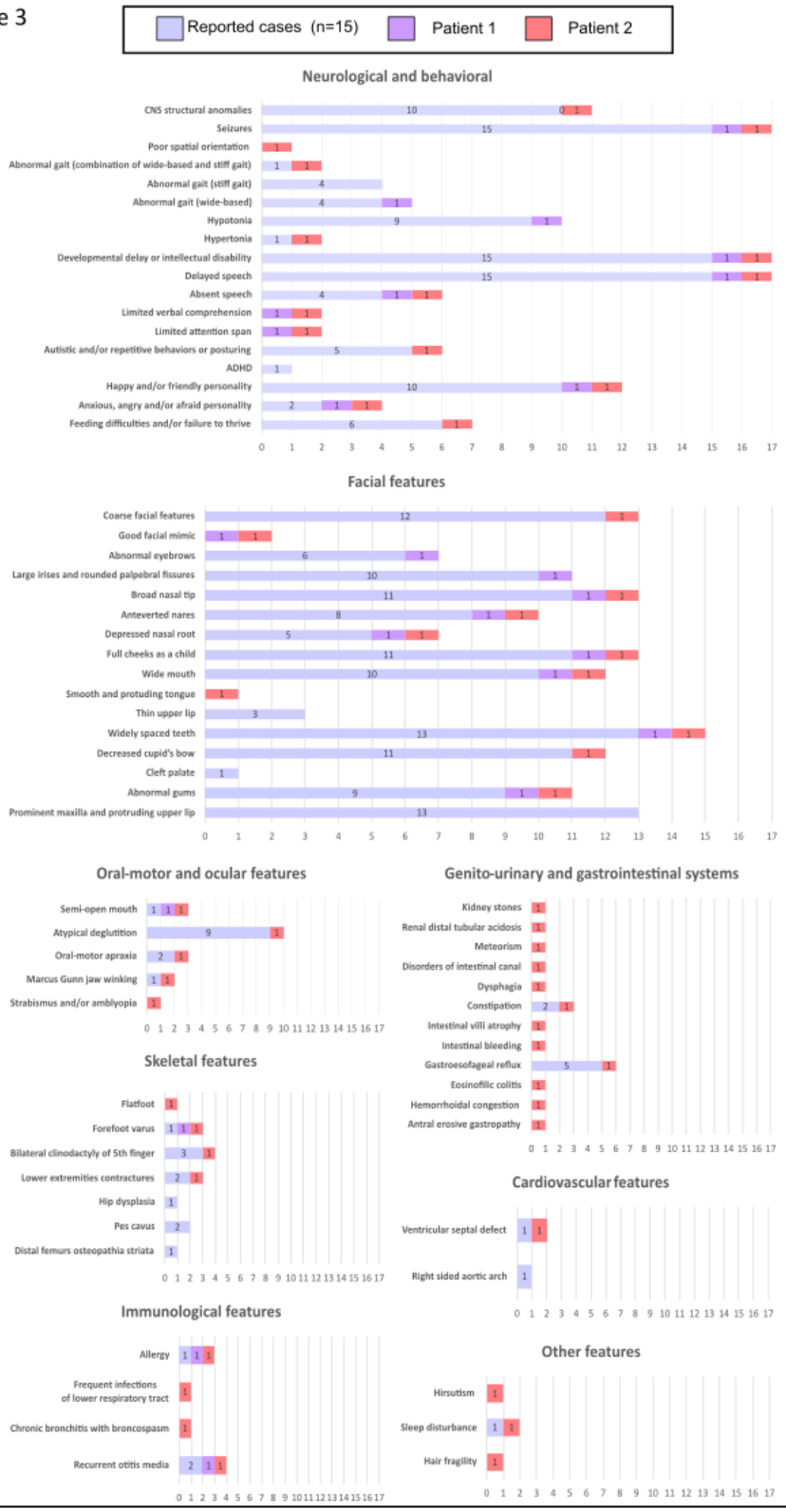

A
POST NATAL DAY 13
MEDICAL REPORT
Cardiovascular defects
3 YEARS OLD
MEDICAL REPORT
Severe language impairment
Impairment of global and motor abilities
Neuropsychiatric impairment
Periventricular leukomalacia
Generalized epilepsy
Renal tubular acidosis
DRUGS
Phenobarbital
Valproate
Sodium bicarbonate
15 YEARS OLD
MEDICAL REPORT
Gastoesophageal reflux
DRUGS
Omeprazole
21 YEARS OLD
MEDICAL REPORT
Drug-resistant epilepsy
Amenorrhea and spaniomenorrhea
Small uterus, Nabothian cyst and multifollicaolar ovaries
PERFORMED TEST
EEG: normal
Pituitary gland and cerebral MRI: normal pituitary gland and cerebral hypoxic perinatal pain
Hormonal dosages: high prolactin and low progesterone levels
23+ YEARS OLD
MEDICAL REPORT
Folate deficit
Insomnia
DRUGS
Dydrogesterone
Calcium folinate
Lormetazepam
Mirtazapine
29 years old: molecular diagnosis
25 years old: enrolment in WES project
PERFORMED TEST
Metabolic diseases screening: normal
Cerebral MRI and TC
MEDICAL REPORT
Slight cerebral MRI and TC alterations
2 YEARS OLD
MEDICAL REPORT
Early pubertal development
Advanced bone age
DRUGS
Leuprolide acetate
8 YEARS OLD
MEDICAL REPORT
Gastroesofageal system impairment
Intestinal bleeding and intestinal villi atrophy
Iron deficiency anemia
Chronic bronchitis with bronchospasm
PERFORMED TEST
HLA typing for celiac disease: negative
Gastroscopy: antral erosive gastropathy
Colonscopy: hemorrhoidal congestion
18 YEARS OLD
MEDICAL REPORT
High homocysteine levels
PERFORMED TEST
MTHFR gene sequencing: inherited variant in the gene associated to reduction in folate
DRUGS
Folic acid supplementation
22 YEARS OLD
B
Patient 2
CT scan (2 yrs.)
MRI (21 yrs.)
T1 sagittal
T2 axial flair
T1 sagital

Figure 3


Reported cases (n=15)
Patient 1
Patient 2
Neurological and behavioral
CNS structural anomalies
Seizures
Poor spatial orientation
Abnormal gait (combination of wide-based and stiff gait)
Abnormal gait (stiff gait)
Abnormal gait (wide-based)
Hypotonia
Hypertonia
Developmental delay or intellectual disability
Delayed speech
Absent speech
Limited verbal comprehension
Limited attention span
Autistic and/or repetitive behaviors or posturing
ADHD
Happy and/or friendly personality
Anxious, angry and/or afraid personality
Feeding difficulties and/or failure to thrive
Facial features
Coarse facial features
Good facial mimic
Abnormal eyebrows
Large irises and rounded palpebral fissures
Broad nasal tip
Anteverted nares
Depressed nasal root
Full cheeks as a child
Wide mouth
Smooth and protuding tongue
Thin upper lip
Widely spaced teeth
Decreased cupid's bow
Cleft palate
Abnormal gums
Prominent maxilla and protruding upper lip
Oral-motor and ocular features
Semi-open mouth
Atypical deglutition
Oral-motor apraxia
Marcus Gunn jaw winking
Strabismus and/or amblyopia
Genito-urinary and gastrointestinal systems
Kidney stones
Renal distal tubular acidosis
Meteorism
Disorders of intestinal canal
Dysphagia
Constipation
Intestinal villi atrophy
Intestinal bleeding
Gastroesofageal reflux
Eosinofilic colitis
Hemorrhoidal congestion
Antral erosive gastropathy
Skeletal features
Flatfoot
Forefoot varus
Bilateral clinodactyly of 5th finger
Lower extremities contractures
Hip dysplasia
Pes cavus
Distal femurs osteopathia striata
Cardiovascular features
Ventricular septal defect
Right sided aortic arch
Immunological features
Allergy
Frequent infections of lower respiratory tract
Chronic bronchitis with broncospasm
Recurrent otitis media
Other features
Hirsutism
Sleep disturbance
Hair fragility

4.2 KCNK18: biallelic variants altering TRESK channel activity are associated with intellectual disability and neurodevelopmental disorders

KCNK18 encodes the TWIK-Related Spinal cord K^+ channel (TRESK), a member of the two-pore domain (K2P) potassium channel family. K2P channels count 15 members, each one encoded by different *KCNK* genes[216]. At a protein level, these channels have different names according to their structural and functional similarities. They are classified into the TWIK, TREK, TASK, TALK, THIK, and TRESK subfamilies[217].

The molecular architecture of the TRESK channel resembles other K2P channel subtypes except that it has a longer intracellular loop between the second and the third transmembrane domains and a shorter C-terminal tail. The intracellular loop contains fundamental modulatory sites such as phosphorylation, dephosphorylation and calcineurin binding motifs. Under resting conditions, murine TRESK is phosphorylated and maintains an inactive conformation but after a calcium signal, the channel is rapidly activated following dephosphorylation by the calcium-dependent phosphatase calcineurin[218,219]. To allow this process, calcineurin must bind to the channel through interaction with its non-catalytic sites, helped by the regulatory protein, 14-3-3. In humans, the process is also supported by protein kinase C (PKC)[218].The switch between the active and resting states has been intensely investigated[219] and it is mediated by the phosphorylation of TRESK at two different regulatory regions, the 14-3-3 binding site (Serine 252, Ser252, hereafter) and three adjacent serine residues (Ser262, Ser264 and Ser267, collectively called serine cluster, hereafter). These residues are also conserved in murine TRESK as Ser264, Ser274, Ser276 and Ser279.

The activated calmodulin complex stimulates calcineurin binding to the TRESK intracellular loop, promoting dephosphorylation of the two regulatory residues (Ser252 and the serine cluster) and the activation of the channel. After the decay of the calcium signal, channel activity is inhibited by restoring of the phosphorylated state by protein kinase A, which acts at the level of Ser252, and microtubule affinity-regulating (MARK) kinases, which mediate phosphorylation at the level of the serine cluster, supported by 14-3-3 proteins[219].

Interestingly, a high basal activity and reduced response to calcium have been reported in a study using a murine TRESK mutant characterized by a constitutively dephosphorylated state of the serine cluster[219]. Remarkably, prolonged phosphorylation of this motif has been shown to cause reduced responsiveness of the channel and low basal activity[219,220], further documenting the complex modulatory process controlling TRESK function.

KCNK18 expression in dorsal root (DRG) and trigeminal (TG) ganglia[221] led to a proposed role for this channel in a variety of pain pathways[222]; moreover, an involvement in typical migraine with aura has been suggested by different authors[223–226]. Despite the expression of *KCNK18* in basal ganglia, its mRNA is detectable also in other brain compartments, including hypothalamus, frontal cortex, cortex, anterior cingulate cortex, hippocampus, spinal cord, and substantia nigra (www.gtexportal.org), suggesting a wider involvement of the channel in physiological brain function. Interestingly, recent studies have also reported *KCNK18* variants in patients with developmental delay (DD) and intellectual disability (ID), with migraine as an associated feature in some cases[227,228]. Here, we describe the identification of a family with three siblings affected by mild to moderate ID, seizures, autistic-like behavior with different degrees of severity, carrying biallelic *KCNK18* variants, supporting a possible involvement of the gene in neurodevelopmental disorders (NDD).

Materials and Methods

Whole exome sequencing, prioritization, and variant calling

DNA was extracted from total blood using the ReliaPrep Blood gDNA Miniprep kit (Promega, Madi-son, WY, USA) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientifics, Waltham, MA, USA). Array-CGH was performed using a 60K whole-genome oligonucleotide microarray (Agilent Technologies, Santa Clara, California, USA).

Patients were enrolled in the Autism Sequencing Consortium (ASC) project and their gDNA samples were sequenced at the Broad Institute on Illumina HiSeq sequencers as previously described[6,26].

ES raw data of the trio were processed and analyzed using an in-house implemented pipeline previously described[27,28], which is based on the GATK Best Practices[29]. The UCSC GRCh37/hg19 version of genome assembly was used as a reference for reads alignment by means of BWA-MEM[30] tool and the subsequent variant calling with HaplotypeCaller (GATK v3.7)[29]. We used SnpEff v.4.3[31] and dbNSFP v.3.5[32] tools for variants functional annotation, including Combined Annotation Dependent Depletion (CADD) v.1.3[33], Mendelian Clinically Applicable Pathogenicity (M-CAP) v.1.0[200] and Intervar v.0.1.6 for functional impact prediction[34]. Thereby, the analysis was narrowed to variants which affect coding sequences or splice site regions. Moreover, high-quality variants were filtered against public databases (dbSNP150 and gnomAD ver.2.0.1) so that only variants with unknown frequency or having MAF <0.1%, as well as variants occurring with frequency <1% in our population-matched database (~2,000 exomes) were considered.

Further variant stratification in conformity with the American College of Medical Genetics and Genomics (ACMG) guideline[35], considering also mode of inheritance and functional in-silico prediction of impact, allowed us considering the final set of variants for possible associations with the phenotype. All variants are referred to GRCh37 annotation and to NM_181840.1.

Identified variants were confirmed by Sanger sequencing using standard conditions and the following primers: 5'-gggagatggcagaaggtctcttta and 5'-ttactcctctccatggcttgtgg; 5'-caaacttggcacatgtccttcac and 5'-gcatgaccctgaaagacaacaca.

Variants were analysed with the VarSome tool[159] as a starting point for further analysis. This allowed evaluation of at least nine in silico predictors simultaneously. Variants frequency in the population was evaluated using Genome Aggregation Database (GnomAD) Browser version 2.1.1.

Constructs

Human TRESK was subcloned between the 5' and 3' UTR of the Xenopus β-globin gene in the oocyte expression vector, pFAW. The identified variants were introduced by site directed mutagenesis and confirmed by automated sequencing. mRNA for wild-type and mutant channels was synthesized using the T7 mMESSAGE mMACHINE kit (Ambion, Life technologies, Carlsbad, CA, USA) and mRNA concentrations were quantified by spectrophotometric analysis prior to injection.

Electrophysiological recordings

Xenopus laevis oocytes collection and defolliculation have been performed according to standard protocols which fall under the international standards of animal care, the Maltese Animal Welfare Act and the NIH Guide for the Care and Use of Laboratory Animals. TRESK wild-type or mutant mRNAs were microinjected into oocytes in equal quantities, unless otherwise stated. Forty-eight hours after injection, whole-cell currents were recorded at room temperature (20–22°C) by using the two-electrode voltage clamp method (Axoclamp-2B, Axon DIGIDATA 1550B, Axon Instruments, Foster City, CA, Axon). Basal K^+ currents were measured in a low K^+ extracellular solution (in mM: 95.4 NaCl, 2 KCl, 1.8 CaCl2, 5 HEPES pH 7.5 with NaOH). Recording electrodes (0.3÷1 MΩ) were backfilled with 3M KCl. Currents were filtered at 100 Hz and digitized at 1 kHz for analysis. From a holding potential of -80 mV, 1-second-long voltage commands were applied from -120 mV to +60 mV, delivered in 20 mV increments. Currents were recorded using Clampex 10.7 software (Axon Instruments, Foster City, CA). Ionomycin induced TRESK currents were measured in high K^+ solution (in mM: 17.4 NaCl, 80 KCl, 1.8 CaCl2, 5 HEPES, pH 7.5 with NaOH) at the end of 300 ms-long voltage steps from a holding potential of 0 mV to -100 mV. Ionomycin (free acid form) was made as a stock solution of 1 mM in DMSO and diluted

in the high K^+ solution to 0.5 µM on the day of the experiment. Data were analyzed with Clampfit 10.7 (Axon instruments, Foster City, CA) and Igor programs.

Statistics

Data are given as mean values ± standard error of the mean (SEM), where n represents the number of oocytes. Results were reproducible in at least 2–3 different batches of oocytes. Statistical significance was determined using a student's t-test. When error bars are not shown they are smaller than the size of the symbol.

Results

Clinical features of the three affected siblings

A family of European origins with three siblings affected by an unclassified NDD (Figure 1A). The parents had no history of NDD or notable diseases, except for a carcinoma in the mother. The first proband was a 26-year-old female at the time of the last evaluation. Born at term after cesarean delivery, the proband showed a delay in the acquisition of developmental milestones, with first phonemes at 18 months, walking at 24 months and sphincter control at 5 years. Mild ID was reported (total IQ 64, verbal IQ 60, performance IQ 70), with difficulties in calculation and inability to recognize the value of money. Poor language, dysarthria and attention deficit were observed. Moreover, partial temporal-spatial disorientation, poor criticism, and judgment with limitation of the possibilities of self-determination were reported. The support from a caretaker was required for everyday activity and for self-care. She suffered frequent (once a week) frontal cephalalgy. At clinical examination, the behaviour was uncooperative and restless. Even if autism spectrum disorder was not formally assessed using Autism Diagnostic Observation Schedule (ADOS-2), nor other behavioural tests, the patient showed an important impairment of relationships that, combined with the other symptoms, led the physician to include her in the autism spectrum.

The second born proband was a 19-year-old female diagnosed with a mild ID (total IQ 56, verbal IQ 47, performance IQ 76). The patient showed compromised social relations that, as for her sister, included the patient into a clinical diagnosis of autism spectrum disorder. Specific difficulties in the logical-mathematical area were reported, together with the inability in remembering new concepts. The support from a second person was necessary for everyday activity. An episode of epileptic seizure was reported and supported by electroencephalography (EEG), that showed comitiality.

The last born was a 16-years-old male affected by dyslexia and difficulties in learning, but no further information were available.

The social-economic situation of the family was also considered, to understand if the impaired behaviour of the patients could be caused by a combination of genetic and environmental factors[229,230]: no noteworthy situation was identified, except for a worsening of the already low social behaviour of the patients after the death of their mother.

Genomic and structural analyses

Patients were enrolled in the ASC project[26], and ES of the whole family members performed. In the three affected members, the analysis led to the identification of compound heterozygosity for two missense changes (c.487T>G, p.Tyr163Asp, paternally inherited; c.755C>T, p.Ser252Leu, maternally inherited) (Figure 1A). Both variants affected residues that are highly conserved among orthologs (Figure 1B) and were predicted to have a disruptive impact on protein function (Supplementary Table 1). Variant p.(Tyr163Asp) had not previously been reported in GnomAD database (ver 2.1.1) (www.gnomad.broadinstitute.org) while variant p.(Ser252Leu) had an allelic frequency of 0.0001379 at the time of the analysis; both the variants were not previously reported in our in-house databases. Both affected residues are located within the intracellular domains of the channel (Figure 1C), a region that plays a key role in the conformational switching events controlling TRESK function[220]. Of note, Ser252 plays an important regulatory role in the switching process by cycling between its phosphorylated and dephosphorylated states[220]. The highly conserved nature of the two residues and their predicted location consistently supported the functional relevance of the two amino acid substitutions. ES data analysis excluded the occurrence of other functionally relevant variants compatible with known Mendelian disorders based on the expected inheritance model and clinical presentation, further supporting the relevance of the two missense variants involving *KCNK18*.

p.Tyr163Asp and p.Ser252Leu do not alter basal TRESK function Genomic and structural analyses.

To validate the *in silico* predictions, we next examined whether the two missense changes in *KCNK18* affect the functional properties of TRESK. Figure 2A shows representative whole-cell basal currents recorded from *Xenopus* oocytes injected with mRNA encoding human wild-type (WT) or the mutant TRESK channels. All channel types exhibited similar outwardly rectifying K^+ current amplitudes although current in oocytes expressing $TRESK^{S252L}$ was slightly larger (Figure 2A, B). This, not statistically different larger basal K^+ current could be due to a reduction in the phosphorylation state of the mutant channel compared to $TRESK^{WT}$. However, a robust gain of function in the basal K^+ current, has been observed when a serine close to the cluster was

replaced with an alanine. Given that the K2P channels work as dimers, to mimic the compound heterozygous condition of affected subjects, equal amounts of the *KCNK18* mutated mRNAs were co-injected into oocytes in a 1:1 ratio. Co-injection of mutant mRNAs did not substantially affect the whole-cell basal currents, which had similar amplitudes to those reported in oocytes overexpressing WT TRESK (Figure 2A, B). TRESK channels typically produce a leak K^+ current that stabilizes the resting membrane potential[228]. We observed that the expression of WT TRESK or both mutants, expressed separately or together, shifted the cell resting membrane potential of the oocytes towards the K^+ equilibrium potential (E_K) (Figure 2C). Assuming an intracellular K^+ concentration of approximately 140 mM in the oocyte, the predicted value for E_K is -107 mV. A similar hyperpolarizing effect was observed upon expression of all TRESK channel types (Figure 2C). Collectively, these findings indicate that both mutations do not dramatically change the basal activity of TRESK channels, which also retain their ability to control the cell resting potential.

<u>p.Tyr163Asp and p.Ser252Leu impair the ability of TRESK to properly respond to calcineurin activation.</u>

Given that the variants lie in amino acids located in regions important for the calcineurin-dependent gating of the channel, we next examined whether the mutant channels retain the ability to respond to calcineurin activation. In WT TRESK channels expressed in *Xenopus* oocytes, an increase of intracellular Ca^{2+} produced by the calcium ionophore ionomycin, generally result in strong calcineurin-dependent activation of TRESK [220,231]. We therefore determined the effect of ionomycin on WT TRESK and the identified variants. As expected, 0.5 µM ionomycin induced a large, robust, and reversible activation of WT TRESK (Figure 3A). By contrast, we found that ionomycin was much less effective in the activation of homomeric and heteromeric mutant channels (Figure 3A). Indeed, the absolute values of Ca^{2+}-activated whole-cell currents were significantly smaller than that for the WT channel (Figure 3B; WT *vs* all groups $p<0.001$). Moreover, we found that the extent of activation (I_{Iono}/I_{Basal} ratios) for both homomeric and heteromeric TRESK channels was much smaller compared to what observed for the WT channel (Figure 3C). However, due to the relatively small amplitude of the unstimulated I_{Basal} currents for the $TRESK^{Y163D}$ and $TRESK^{S252L}$ mutants these values may be overestimated. The response of un-injected control oocytes to ionomycin was insignificant, as previously reported [228,232] (Figure 3A-C) and all groups showed an obvious difference in the ionomycin-induced current when compared to un-injected oocytes (Figure 3B; WT *vs* un-injected $p<0.001$; Y163D *vs* un-injected $p=0.02$; S252L and Y163D+S252L *vs* un-injected $p<0.01$). The relatively small ionomycin-induced

current amplitude recorded from mutants, from Y163D, could be responsible for the misleading reduced level of significance when compared with that of control un-injected oocytes.
Overall, these results indicate that although the identified disease-associated *KCNK18* variants do not alter the basal activity of TRESK, both markedly impair the ability of the channel to respond to calcineurin activation.

Discussion

TRESK background K^+ channel contributes to the stabilization of the resting membrane potential of sensory neurons, particularly at the level of dorsal root and trigeminal ganglia[233,234]. Variants in *KCNK18* gene have been frequently associated with migraine[224,226,228,235]. Interestingly, patients with ID, with or without migraine, have also been reported [227,228], suggesting a contribution of dysregulated TRESK function in NDDs.
In the frame of a large genome scan project focused on patients with NDDs[26], we identified a family of European origin with three siblings affected by ID. In the oldest case (patient 1), recurrent episodes of migraine were also reported. In this family, trio-ES allowed the identification of a new compound heterozygosity for two missense variants (Tyr163Asp and Ser252Leu) in all affected siblings. The variant occurring on residue Ser252 caught our attention, as it has been largely demonstrated that this residue is fundamental for the regulation of TRESK, being subjected to reversible phosphorylation, regulating the activation of the channel[219,236,237]. In pharmacophore-based virtual screening to reveal novel inhibitors against migraine, it has also been suggested Ser252 is one of the binding sites for Ergotamine and Lasmiditan, two drugs commonly used for migraine treatment[238].
To functionally validate the pathogenic relevance of the identified amino acid substitutions on TRESK function, electrophysiological recordings using *Xenopus* oocytes were performed. Slightly larger whole-cell basal K^+ current amplitudes were observed in oocytes expressing $TRESK^{S252L}$, that could be attributed to a reduction in the phosphorylation state of the mutant channel compared to $TRESK^{WT}$. Indeed, a similar effect have been reported when the serine in position 276 ($TRESK^{S276A}$) in the rat clone, was replaced with an alanine[239]. On the other hand, the substitution of serine 276 with glutamate, mimicking the dephosphorylated state, resulted in a mutant with low basal activity and reduced responsiveness. Interestingly, $TRESK^{S276A}$ also shows a reduced response to calcium signals. The S252 residue is part of the binding motif for 14-3-3, a family of ubiquitous adapter/regulatory proteins involved in the phosphorylation/dephosphorylation-dependent regulation of the channel. It has been shown

that the 14-3-3γ isoform directly binds to the intracellular loop of TRESK and control the kinetics of the calcium-dependent regulation of the channel[240].

Notably, 14-3-3γ plays a role in neuronal migration, morphology, and brain development[241]. The gene encoding this protein (*YWHAG*), has been associated to developmental and epileptic encephalopathy (OMIM #717665), and autistic traits have been described in affected patients[242]. It is therefore intriguing to hypothesize that S252L variant could cause an alteration in the interaction with 14-3-3γ, impairing its downstream activity and altering the same pathways involved in *YWHAG*-related disorder[243].

Here, we demonstrated that both Y163D and S252L variants remarkably affect TRESK response to intracellular Ca^{2+}, which can cause calcineurin-dependent activation of the channel[220,231]. Specifically, while ionomycin induced a large and reversible activation of the WT channel, this activation was strongly reduced by the two variants, in both homomeric and heteromeric mutant channels. The exact location of both variants in TRESK channel cannot be determined due to the lack of any 3D structural information for the regulatory domain of TRESK channels. However, in the linear sequence, Y163 is located near the calcineurin binding site. It could therefore be hypothesized that its substitution with aspartate (Y163D), a negatively charged residue, could disrupt calcineurin-channel interactions, and abolish current activation (Figure 3). It is to note that the impairment of the channel activation after ionomycin induction (Figure 3) is not representative of the effect of the variants in the parents, as both the constructs were not expressed in combination with the WT one. With these experiments, we are therefore mimicking more likely a homozygous state of the variants instead of the heterozygous state observed in the parents. This is also in line with the higher impairment of the channel activation observed with Y163D and S252L alone compared to their combination.

The TRESK subfamily consists of only one member (*KCNK18*, K2P18.1) that has been cloned from human spinal cord[244]. TRESK was shown to be expressed in rodent cerebrum, cerebellum, brain stem and, spinal cord[220]. Published evidence and data retrieved from Allen Brain Atlas (www.portal.brain-map.org/) and GTEx Portal (www.gtexportal.org) clearly show high expression of TRESK channels in human brain regions, some for which the impairment has been associated with autism and ID[245]. How mutations in TRESK channels identified so far[227,228] alter the functionality of these brain areas and result in autism and ID remains to be investigated. Whatever the final impact on the human brain of the mutations described here, which disrupt the calcineurin-dependent regulation of TRESK may be, some insights could be gained by considering the effects of calcineurin inhibitors.

Indeed, Cyclosporin A and FK506 are inhibitors of the calcium/calmodulin-dependent protein phosphatase calcineurin, which abolish calcium-dependent TRESK activation. Neurological complications of Cyclosporin therapy are frequent and include dysarthria, headache, restlessness, and seizures. Neurologic toxicity related to therapy with Tacrolimus (FK506), a macrolide calcineurin inhibitor, has also been reported, including restlessness, agitation, confusion, sleep problems, mental depression, dysarthria, headache, and seizures. Intriguingly, dysarthria, restlessness, headache, and seizures have been observed in our patients and could be tempting to speculate that these symptoms could be associated, at least in part, with Y163D- and S252L-induced disruption of calcineurin-dependent TRESK channel activation in the brain. Recently, it has been proposed that calcineurin regulation of two-pore potassium-leak channel activity (*Kcnk5b*) mediates activation of developmental gene transcription that control a broad number of developmental pathways[246]. Similarly to TRESK channels, inwardly rectifying K^+ channels (Kir) control the resting potential of neurons and K^+ homeostasis in the brain. Previous findings showed the importance of several Kir channel types in neurodevelopment, autism, and ID[247–253]. Thus, growing evidence shows that K^+ dependent bioelectric mechanisms can regulate not simply the excitability of neurons but also cell migration, proliferation, differentiation, gene transcription, promoting coordinated developmental signaling.

In summary, our study suggests that the *KCNK18* variants observed in our patients dramatically impair the ability of TRESK channel to respond to calcineurin activation, leading to alteration of a crucial function of the channel. Indeed, acetylcholine, glutamate, or histamine activate native TRESK currents through G-protein coupled receptors[254] and this function would be nearly abolished by both Y163D and S252L variants. Collectively, our findings strengthen the notion that abnormal TRESK channel function in the central and peripheral nervous system could play crucial roles in neurologic and psychiatric diseases. Further studies on larger cohorts of patients and the characterization of transgenic animal models of the disease are however required to elucidate the link between migraine, autism, epilepsy, and intellectual disability observed in some patients carrying TRESK variants.

Figures

Figure 1. Pedigree, conservation, and graphical representation of the variants.

(A) Pedigree of the family and Sanger sequencing chromatograms documenting co-segregation between compound heterozygosity for the two missense changes in *KCNK18* and clinical trait transmitted in the family. **(B)** Multiple alignments of the amino acid stretches flanking the affected residues in TRESK orthologs showing conservation of both mutated residues (*in red*).

Residues are referred to human *KCNK18*, NM_181840.1. **(C)** Schematic topology of the human K2P TRESK channel, showing location of Tyr163 (Y163) and Ser252 (S252) (red asterisks). Human TRESK is composed of 384 amino acid residues and shows the most characteristic features of K2P channels, including four transmembrane helices, two pore-forming domains, and intracellular N- and C-termini. A peculiar structural feature of TRESK (compared to other K2P channels) is the long intracellular loop between the second and third transmembrane segments (TMS), and the relatively short C-terminal tail[218]. In the image, representative domains are shown: PQIIIS and LQLP are calcineurin-binding motifs[220,255], while the serine cluster and Ser252 are indicated by the letter S and are respectively MARK kinases and 14-3-3 and PKA binding motifs.

Figure 2. p.Tyr163Asp and p.Ser252Leu do not significantly affect the basal current and ability of TRESK channels to control the cell resting potential.

(A) Representative families of current traces for the indicated channel types. Cell membrane potential was held at -80mV. Each family of currents was evoked by voltage steps from -120 mV to 60 mV, with 20 mV increments. The recordings were performed 48 hours after the injection of 1ng of mRNA for each channel type. (**B)** Average IV relationships calculated from experiments as in A for the labelled channel types (color coded). The data points are mean ± standard error (*** p<0.001; n=12).(**C)** Resting membrane potentials from individual oocytes un-injected (grey dots), injected with 1 ng of WT (black dots) or Y163D (green dots) or S252L (blue dots) or 0.5 ng of each Y163D and S252L (red dots) mRNA and recorded 48 hours after injection. Note that the expression of all channel types shifts similarly to the resting potential toward K^+ reversal potential. The dots in C represent single cell recordings. Numbers are mean ± standard deviation.

Figure 3. *KCNK18* p.Tyr163Asp and p.Ser252Leu impair the ionomycin-induced activation of TRESK channels.

Representative data points showing ionomycin activation of TRESK WT, Y163D, S252L and Y163D+S252L currents as well as endogenous currents from un-injected oocytes. (**A)** Currents were evoked by 300 ms long voltage steps from 0 mV to -100 mV from oocytes injected with 1 ng of mRNA of the corresponding and indicated homomeric channel type and 0.5 ng for each subunit for the heteromeric channel Y163D/S252L. The sampled data represent the average of 50 ms long period of the steady-state currents recorded at -100 mV. Ionomycin activates WT channels but has little effect on the oocytes expressing the mutant channels or un-injected. Note that the current decay upon reapplication of a solution containing 2 mM K^+ in the recording chamber is due to the reduced K^+ concentration rather than ionomycin washout. (**B)** Bar graph showing the mean of ionomycin activated currents (*** p<0.001). All the groups were

statistically significant compared to uninjected group (WT *vs* uninjected p<0.001; Y163D *vs* uninjected p=0.02; S252L and Y163D+S252L *vs* uninjected p<0.01; n=12). (**C)** Each point represents the steady state ionomycin activated current divided by the steady-state basal current (I_{Iono}/I_{Basal}) recorded in the presence of 80 mM extracellular K^+.

Figure 1

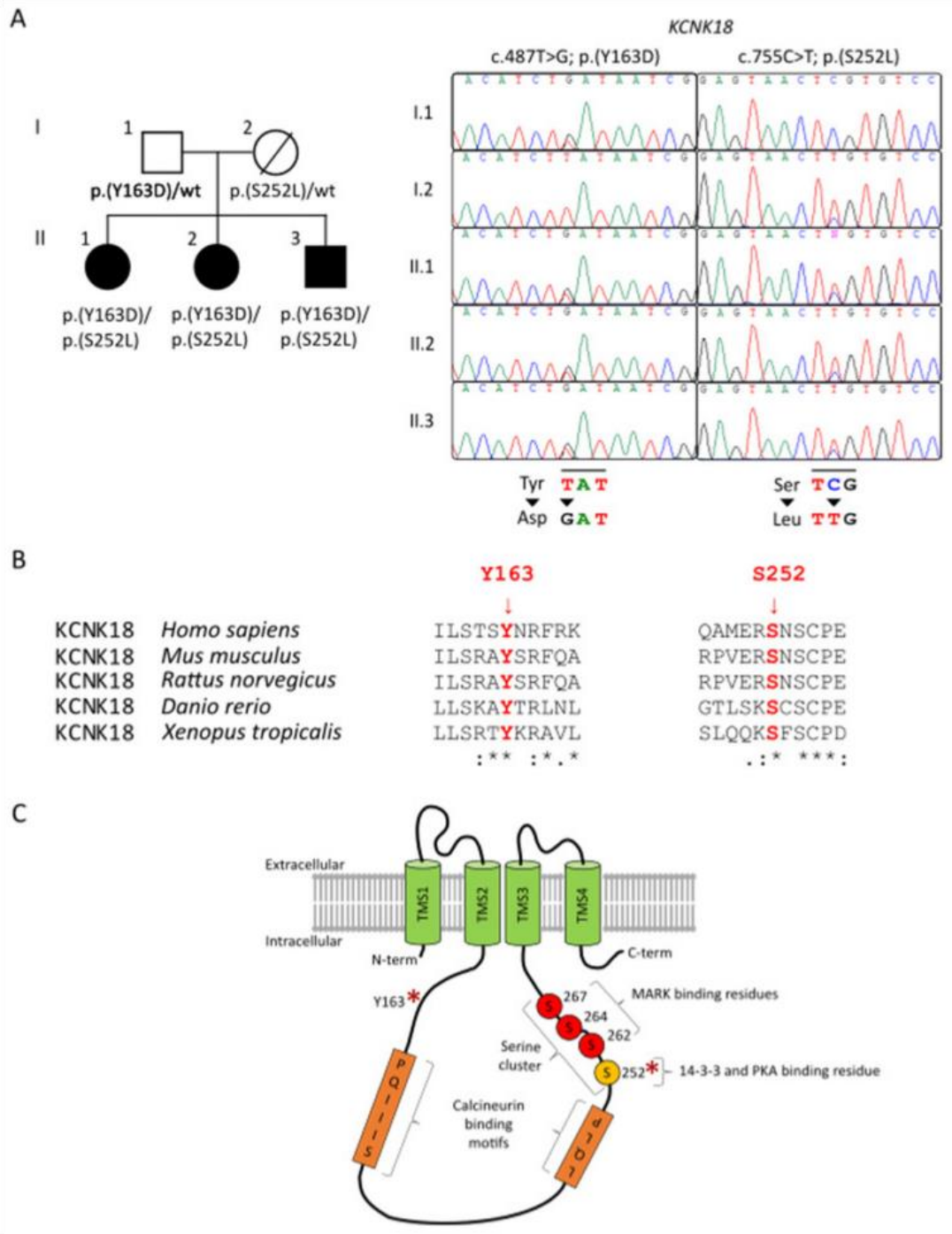


Figure 2

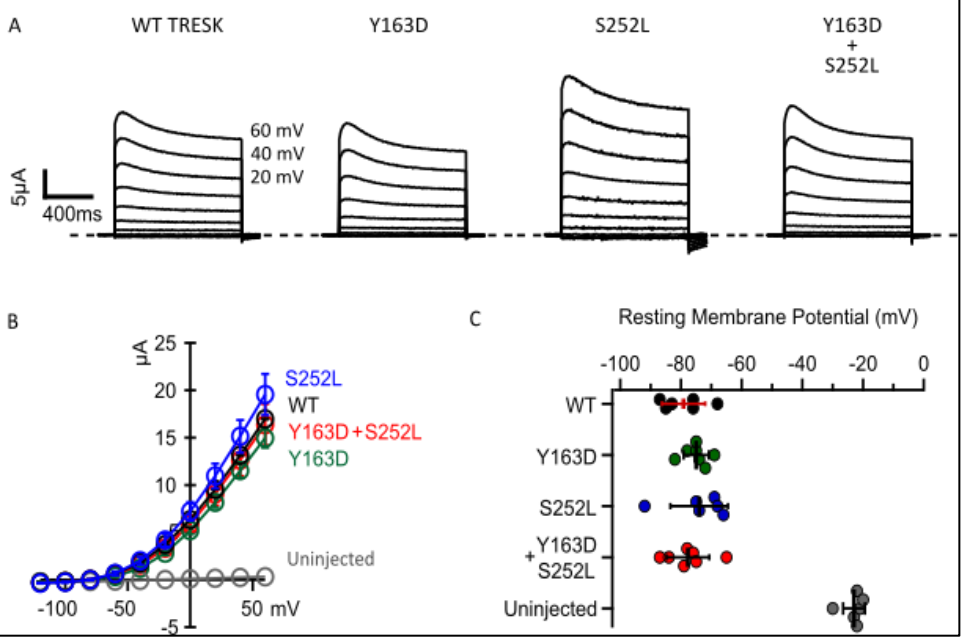
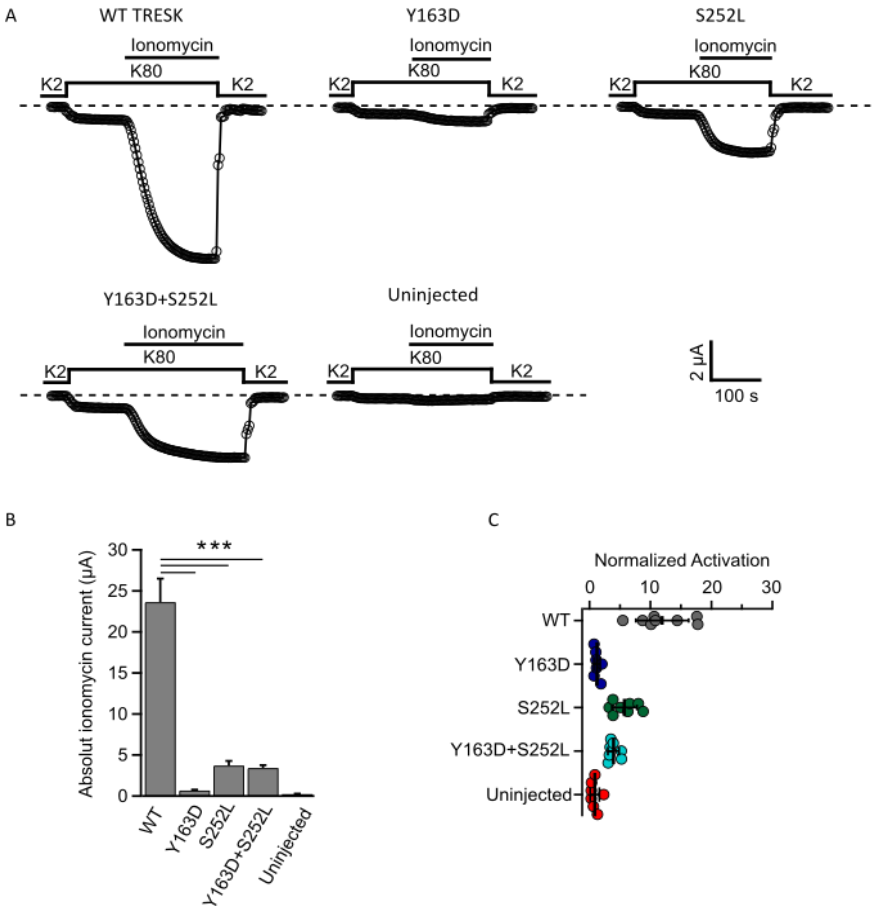


Figure 3



4.3 SUPT5H variants are associated with a NDD with language impairment, motor delay and dysmorphic features

Transcription by RNA polymerase II (Pol II) is a high regulated mechanism in mammals, involving many factors controlling each transcriptional step[256–259] . Among them, Spt5 (encoded by *SUPT5H* gene), is involved in most of the transcriptional steps, and it is one of the most highly conserved transcription regulators, sharing sequence similarities with bacteria, archaea, yeast and mammals[260].

It has been debated about eventual tissue-specific functions for Spt5 as, for instance, it is required for sustaining most Pol II-mediated sense gene transcription in *Drosophila* and *S. pombe*, but it has been shown to regulate only a group of genes in zebrafish, mouse and human cells[260]. However, recent studies[260], demonstrated Spt5 is a crucial positive regulator of global transcription in human cells and loss of Spt5 has been shown to cause reduced expression of most of its target genes. In addition, it has been shown an essential role in maintaining cellular and chromatin Pol II levels, and an acute depletion of Spt5 causes a global and persistent defect in the processivity of Pol II in human cells[260].

To date, *SUPT5H* gene is still not associated to any disease in OMIM database (www.omim.org). However, variants in this gene have been reported as associated with a β-thalassemia phenotype, with an increase in Hb A_2, in 29 patients from 12 unrelated families[192,261] . Eight out of 12 reported variants were nonsense, one was a small deletion and three were splicing variants.

In our study, we report 12 novel patients with heterozygous variants in *SUPT5H* and a neurodevelopmental phenotype variably associated with elevated Hb A_2, proposing a novel disorder associated with this gene.

Results

Identification of patients with *SUPT5H* variants

In our cohort, we found a patient with a *de novo* splicing variant *SUPT5H* (NM_003169.3) c.320-1G>A; p.(?), with mild intellectual disability, language impairment, facial dysmorphisms, heteroaggressivity, emotional instability and impaired motor skills. Thanks to GeneMatcher, we found additional eleven patients. Six were female and six were male, ranging from 13 months to 47 years of age (Table 1). One patient had a small deletion, two a splicing variants and nine had a missense variant, of which two of them - p.(Arg334Trp) and p.(Arg365Cys) - were found each one in two unrelated cases (Fig. 1A). 12/12 patients showed language impairment, motor delay and intellectual disability, of which half of the patients had mild/moderate ID and the

other half severe ID. 6/12 cases showed epilepsy, 5/9 ASD, 3/12 axial or generalized hypotonia and 2/12 limb hypertonia or spasticity. EEG anomalies were reported in 5 out of 7 tested patients and all of them were the ones for which epilepsy was observed. Additional features were facial dysmorphisms (9/12) (Fig. 1B), frequent respiratory tract infections (3/12) and increased Hb A_2 (3/6). A complicated pregnancy was reported for 5/11 patients.

Supt5h$^{+/-}$ mouse model shows anomalies of the corpus callosum and the primary motor cortex

A heterozygous knock-out mouse model has been generated by Dr Binnaz Yalcin (INSERM, France), and an evaluation of gross defects did not show any peculiar phenotype. On the contrary, using a quantification approach of 66 morphological and 115 cellular measurements across 19 different brain regions in two histological sections at Bregma +0.98 mm and –1.34 mm, we found that heterozygous lack of Supt5h was associated with neuroanatomical phenotypes (Fig. 2A-B). An overall slight reduction of the total brain area, cingulate cortex area, width of the cingulate cortex in the left hemisphere and height of the primary motor cortex in the right hemisphere were noticed (p value 0.05) in the striatum section (Bregma +0.98 mm). In addition, a significant reduction in the height of the primary motor cortex of the left hemisphere (p value $1x10^{-2}$) and in the width and height of the genu of the corpus callosum (p value $1x10^{-3}$) were observed in the same area. At the level of the hippocampus section (Bregma -1.34 mm), we observed a non-significant reduction of the height of the secondary somatosensory cortex of the right hemisphere and of the left habenula and a slight reduction (p value 0.05) of the total brain area, the dorsal hippocampal commissure, the secondary somatosensory cortex of the left hemisphere and of the intermediate cortical layer. Furthermore, a significant decrease (p value $1x10^{-3}$) in the height of the primary motor cortex was documented (Fig. 2A-B).

A review of all the available brain MRI (7/12 cases) showed that all the analysed parameters were in the range of normal, however below the median according to the age of the patients (Fig. 2C and Table 2).

Ongoing studies

The study of a novel gene without any patients' sample available can be difficult, especially when most of the variants are missense ones that require *ad hoc* studies. Due to the role of the gene in transcription elongation and in the regulation of chromatin structure[260,262], we sent five patients from *SUPT5H* cohort (Table 1) to the EpiSign consortium, to detect any eventual episignature associated with variants in this gene (*see previous Chapters*). We are also performing a deep *in silico* characterization for all the variants, in collaboration with Matteo

Rossi Sebastiano (University of Turin). Preliminary results report an overall rigidification of the protein due to the missense changes, with a possible influence on the interaction with binding partners. Finally, we started a collaboration with Face2Gene[111] (www.face2gene.com/), to identify any overlapping facial feature between our patients, thanks to artificial intelligence.

Discussion

The clinical features observed in our patients are largely overlapping between each other, with a marked neurodevelopmental phenotype. The structural brain defects observed in the *Supt5h*$^{+/-}$ mouse model at the level of the cortex and corpus callosum suggest a potential involvement of this gene in a neurological/neurodevelopmental phenotype. Interestingly, similar malformations have been observed also in MRI imaging data from a group of patients, with the size of some brain areas under the median.

In addition, preliminary *in silico* data show a pathological impact of the missense variants identified, with a modest rigidification of the protein structure and a potential alteration of Spt5 interactome.

Much work is still ongoing, and potential important data will derive from the analysis of the methylation profile, looking for a potential episignature for *SUPT5H* patients. If an episignature will be identified, the next steps would be to increase the number of the patients and to compare the cohort of patients with a neurodevelopmental disorder to the ones with β-thalassemia trait only.

A tricky point of the study is the difficulty to explain which kind of mutations are causative for the β-thalassemia trait, which ones for the neurodevelopmental phenotype and which ones for both. An observation is that we did not identify any frameshift or stop mutation in our cohort, as instead was reported for the β-thalassemia trait. We however identified two splicing mutations, for which the effect must be well characterized. Based on *in silico* predictions, the variant c.524+1G>A; p.(?) (case #S02) should cause an in-frame skipping of exon 7, likely escaping mechanisms of nonsense mediated decay. Splicing analysis on patient-derived PBMC of the variant c.320-1G>A; p.(?) (case #S01) revealed a partial intron retaining of 52 bases of intron 5, with the expression of both mutant and wild-type allele (Figure 3). It remains to be understood how this variation impact on the protein structure and stability.

The partial loss of *Supt5h* in the mouse model causes a phenotype in line with the one observed in our patients and therefore we cannot exclude the hypothesis of a loss of function mechanism. Also, it must be noted that the gene is intolerant to haploinsufficiency (GnomAD pLi score 1) and to missense variation (GnomAD Z score 4.56), and it is therefore plausible to hypothesize that

its disruption should cause important phenotypic effects. Indeed, genetic variants underlying life-threatening diseases, are gradually and selectively eliminated from the population through negative selection[263]: the pLi and Z score calculated for *SUPT5H*, suggest the gene is likely candidate to be linked to a severe condition, and we hypothesize the β-thalassemia trait previously reported is rather a sub-phenotype of the neurodevelopmental disorder, rather than the main feature associated with the gene. Further studies comparing the different cohort of cases will be the key to better understand the pathophysiological mechanisms involving *SUPT5H*, and a key point will be to obtain a deep clinical characterization for both ours and already published cases.

Figures and tables

Table 1. Genetics and clinical features of *SUPT5H* patients

	M01	M02	M03	M04	M05	M06	M07	M08	M09	S01	S02	D01
Genetics												
cDNA	c.370C>T	c.527°>G	c.652C>G	c.1000C>T	c.1000C>T	c.1093C>T	c.1093C>T	c.2486C>T	c.3196C>T	c.320-1G>A	c.524+1G>A	c.1301_1303del
Coding Effect	p.(Arg124Cys)	p.(Asp176Gly)	p.(Pro218Ala)	p.(Arg334Trp)	p.(Arg334Trp)	p.(Arg365Cys)	p.(Arg365Cys)	p.(Pro829Leu)	p.(Arg1066Cys)	p.(?)	p.(?)	p.(Ile434del)
Inheritance mode	*d.n.*	*d.n.*	*d.n.*	*d.n.*	Not maternal; father n.a.	Not paternal; mother n.a. (egg donor)	*d.n.*	Not maternal; father n.a.	*d.n.*	*d.n.*	*d.n.*	*d.n.*
Gender	M	M	M	M	M	M	F	F	M	F	F	F
Age at last examination (years)	17	27	3	15	4	8.5	47	15	13	10	2.6	1.1
Growth												
Birth weight (kg; percentile; SD)	3.57 (51;+0.02)	1.59 (<1;-3.15)	4.00 (81;+0.89)	2.67 (8;-1.43)	2.80 (11;-1.22)	3.18 (27;-0.61)	n.a.	3.43 (50;+0.01)	4.18 (81;+0.89)	n.a.	3.07 (25;-0.68)	2.13 (1;-2.46)
Birth length (cm; percentile; SD)	50 (48;-0.06)	n.a.	n.a.	n.a.	47 (12;-1.18)	50.5 (55;+0.13)	n.a.	n.a.	52 (76;+0.70)	n.a.	46 (6;-1.55)	43 (<1;-2.89)
Birth Head Circumference	37 (75;+0,68)	n.a.	n.a.	n.a.	33.3 (13;-1.14)	34 (20;-0.83)	n.a.	n.a.	36.2 (78;+0.78)	n.a.	33.5 (19;-0.88)	32 (3;-1.90)

(cm; percentile ; SD)												
Weight (kg; percentile ; SD)	59.2 (34;-0.42)	n.a.	11.7 (3;-1.92)	n.a.	14.7 (17;-0.96)	39 (99;+2.18)	90 (98;+1.97)	50.7 (42;-0.19)	55 (77;+0.73)	n.a.	12.3 (12;-1.16)	7,5 (1;-2.2)
Height (cm; percentile ; SD)	183 (91;+1.35)	192 (<1;-21.89)	83 (<1;-3.42)	175 (76;+0.70)	89.5 (<1;-3.04)	135 (89;+1.22)	n.a.	153.9cm	165 (87;+1.13)	n.a.	87.0 (2;-1.96)	70,5 (13;-1.11)
Head Circumference (cm; percentile ; SD)	54.5 (34;-0.40)	58 (98;+2.02)	51 (81;+0.86)	n.a.	n.a.	55 (98;+2.03)	56.7 (98;+2.15)	52.5cm	n.a.	n.a.	43.7 (<1:-3.07)	45 (49;-0.03)
Neurological/Developmental												
Speech/Language	Delay	Delay	Delay	Delay	Delay	Delay	Delay	Delay	Regression from 17 months	N.a.	Absent	Delay
Motor milestones	Delay	Delay	Delay; axial hypotonia	Delay; axial hypotonia and spasticity of the legs	Delay	Delay	Delay; hypotonia	Balance issues	Toe walker	Toe walker; disharmony of perceptual motor integration	No support of the head; profound generalized hypotonia	Delay; severe axial hypotonia and hypertonia of the limbs
Intellectual disability	Moderate	Moderate	Moderate	Severe	Severe	Severe	Severe	Moderate	Moderate	Moderate	Severe	Severe
Seizures / Epilepsy	No	Yes	Yes, myoclonia with fever	Yes	No	No	No	Yes, myoclonic absence seizures and prolonged tonic-clonic seizures.	Yes	No	Yes, febrile	No
Autism	No	Yes	Too young to assess	Yes, severe	No	No, but repetitive behaviour	Yes	Yes	Regression from 17 months	No	Too young to assess	n.a.

EEG	n.a.	n.a.	Normal	Multifocal epileptiform activity	Normal	n.a.	n.a.	Generalised spike, wave/polyspike and wave activity	Active EEG with near continuous DC during sleep	n.a.	Generalized intermittent rhythmic delta slowing; no epileptiform activities	Low amplitude activity, poor physiologic graphoelements; no epileptiform activities; poor spatial organization
Other Behaviour /Psychiatric abnormality	No	Psychotic episodes started at age 20, followed by decline of cognitive and behavoral capabilities	Social-emotional issues	No	No	Opposition, anxiety	Severe ADHD, severe OCD, aggressive behaviour, sleep problems			Eteroaggressivity, emotional instability	Sleep dysregulation	
Other findings												
Hemoglobinopathy (Hb A_2)	n.a.	Normal	n.a.	n.a.	Normal	Normal	n.a.	Normal	Normal	Increased	Increased	Increased
Other findings	Single transverse palm fold	Frequent respiratory tract infections in childhood	Neonatal polycythemia; feeding difficulties in first year (tube feeding); failure-to-thrive; obstipation;		Low T cells; aspiration requiring tube feeds; primary hypothyroidism; lymphatic abnormality with a history of	Right sigle palmar fold; 2nd and 3rd toes syndactyly; iga and igm deficiency	Dental anomalies (missing upper front teeth) and lower teeth severely eroded; small hands				A few hemangioma; restrictive lung disease; multiple respiratory infections); severe obstr	Liver hemangioma; intrapyloric leiomyoma

			frequent respiratory tract infections; eczema.		chylous and pleural effusions; T11 hemivertebrae; Tethered cord		and feet				uctive sleep apnea	
Episign analysis ongoing	DNA n.a.	Yes	Yes	Yes	DNA n.a.	DNA n.a.	DNA n.a.	DNA n.a.	DNA n.a.	Yes	DNA n.a.	Yes

M= male; F= female; n.a.= not available; d.n.= de novo; OCD= obsessive compulsive disorder; ID= intellectual disability; ADHD= attention deficit hyperactivity disorder

Table 2. Review of brain MRI data. For each measurement, the median value adjusted for the age is reported; light orange indicates a value below the median, green indicates a value above the median. Measurements are done according to Garel *et al*. 2011[264]

		Measurements									
Case	**Age at scan**	**APD**	*age: p3-p97 median*	**GT**	*p3-p97 median*	**BT**	*p3-p97 median*	**IT**	*p3-p97 median*	**ST**	*p3-p97 median*
M01	8y	57.2 mm	*8y:57.9-76 M:66.6*	10.4 mm	*7.2-13.1 m:10.3*	5.63 mm	*3.9-7.9 m:5.8*	3.69 mm	*2.2-5.7 m:3.7*	8.54 mm	*7.5-12.8 m:10.1*
M02	18y	67.6 mm	*15y: 62.5-82.3 M:72*	9.3 mm	*7.6-13.7 m:10.8*	5.2 mm	*3.7-8.4 m:5.9*	5.0 mm	*2.5-6.2 M:6.2*	9.2 mm	*6.3-14.8 m:10.5*
M03	11m	52.2 mm	*1y:47.9-66.8 M:55.6*	7.2 mm	*4.6-9.7 m:6.9*	6.2 mm	*2.2-5.7 m:3.6*	4.5 mm	*1.5-4.3 M: 2.5*	6.7 mm	*4.4-9.9 m:6.7*
M05	3y10m	59.5 mm	*3y: 54.4-72.5 M:62.9*	5.04 mm	*6.3-11.8 m: 9.1*	2.6 mm	*3.3-7 m:5*	1.72 mm	*1.8-4.9 m:3.1*	5.17 mm	*6.2-11.5 m:8.8*
S01	3y3m	63 mm	*3y: 54.4-72.5 M:62.9*	6.95 mm	*6.3-11.8 m: 9.1*	4.5 mm	*3.3-7 m:5*	2.3 mm	*1.8-4.9 m:3.1*	8.57 mm	*6.2-11.5 m:8.8*
S02	20m	37.4 mm	*1.5y: 50.6-69.1 M:58.6*	2.02 mm	*5.2-10.4 M: 7.7*	1.71 mm	*2.6-6.1 m:4.1*	1.35 mm	*1.6-4.5 M2.7*	4.63 mm	*5.1-10.5 M:7.5*
D01	1m	41 mm	*0m: 36.8-62 M:43.6*	2.93 mm	*2.5-8.3 m:4.3*	1.82 mm	*1.3-5 m:2.3*	2.17 mm	*1.2-3.9 m:1.9*	3.61 mm	*1.9-9 m:3.9*

CC= corpus callosum; APD= anteroposterior diameter of the CC; GT= thickness of the genu; BT= thickness of the body; IT= thickness of the isthmus; IT= thickness of the isthmus; ST= thickness of the splenium; y=years; m=months

Figure 1. Variants identified in patients and facial features.

(A) Mutation tolerance at each position in Spt5 (*SUPT5H*, NM_003169.3) human protein as calculated by MetaDome. The three main domains of the protein are reported in light purple and are (from left to right) Spt5 transcription elongation factor acidic N-terminal (aminoacids 75-172), Early transcription elongation factor of RNA pol II, NGN section (aminoacids 178-264) and KOW motif (aminoacids 475-501). Variants in codon 334 and 365 have been found in two unrelated patients, as indicated in the figure. Overall, all the variants are located in regions that are intolerant (yellow and orange) or highly intolerant (red) to changes. **(B)** Facial features of some of the patients. (C) Graphical representation of most common features observed in the patients. Bars in magenta indicate the number of cases in which the feature has been observed; the line in purple indicates the total number of evaluated patients.

Figure 2. Cerebral anomalies in *Supt5h*$^{+/-}$ mouse model and in patients

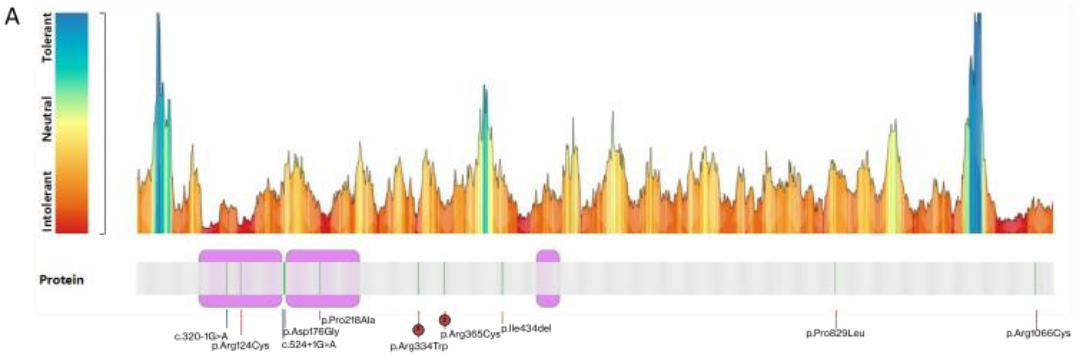
(A) Brain features plotted in two coronal planes (Bregma +0.98 mm and Br -1.34mm) according to P values for *Supt5h* gene. White indicates P > 0.05, and gray indicates no data. Histograms of percentage changes relative to wildtype animals (100%) are colored according to the

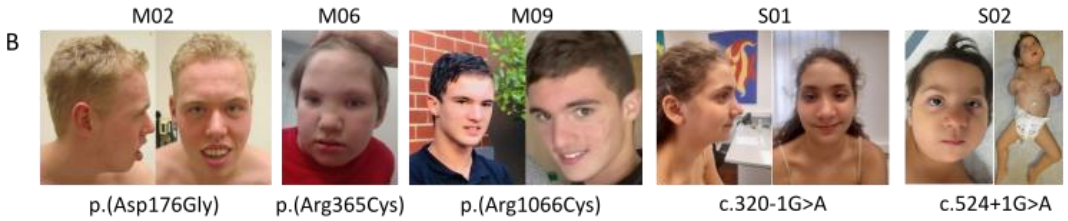
significance level. TBA=Total Brain Area; Cg= Area of the cingulate cortex; Cg_width=Width of the cingulate cortex; gcc=Genu of the Corpus Callosum; M1=Height of the primary motor cortex; dhc= Dorsal hippocampal commissure; S2=height of the secondary somatosensory cortex; ic=area of the internal capsule; hb=Habenula; L=left hemisphere; R=right hemisphere. A complete legenda is reported in the work from Nguyen *et al*.[265]. **(B)** Further insights about measurements of the corpus callosum (left) and primary motor cortex (middle and right) are reported. Brain slices have been stained with cresyl violet and luxol blue and acquired at magnification 20x. **(C)** Representative images of brain MRI of patients described in table 2.

Figure 3. Splicing analysis of c.320-1G>A; p.(?) (case #S01)

(A) Splicing analysis performed on cDNA derived from patient's PBMC. cDNA was amplified using the primers 5'-GACGAGGACCAGTGGGAGGATG and 5'- GATTGGGATCCTTGACTCCTGG (intron spanning) and amplificates were run on an 2% TBE-agarose gel. A shift of less than 100 bp between the wildtype (wt) and mutant (mut) was observed. **(B)** Bands were gel-excised, and the DNA extracted by GenElute Gel Extraction Kit (Sigma-Aldrich). Fragments were sequenced by Sanger method (not shown), showing the retaining of 52 bp of intron 5 (5'-GCCCCTGCCCTCCACCTGCCCTTGCTCACGCCCTCTGCCCATTATTTTTCAA). Sequence from exon 1 to 5 was translated into protein sequence using the tool Expasy (https://web.expasy.org/translate/), showing the insertion of a novel aminoacids sequence and of a stop codon.

Figure 1


A
Tolerant
Neutral
Intolerant
Protein
c.320-1G>A
p.Arg124Cys
p.Asp176Gly
c.524+1G>A
p.Pro218Ala
p.Arg334Trp
p.Arg365Cys
p.Ile434del
p.Pro829Leu
p.Arg1066Cys


B
M02
M06
M09
S01
S02
p.(Asp176Gly)
p.(Arg365Cys)
p.(Arg1066Cys)
c.320-1G>A
c.524+1G>A

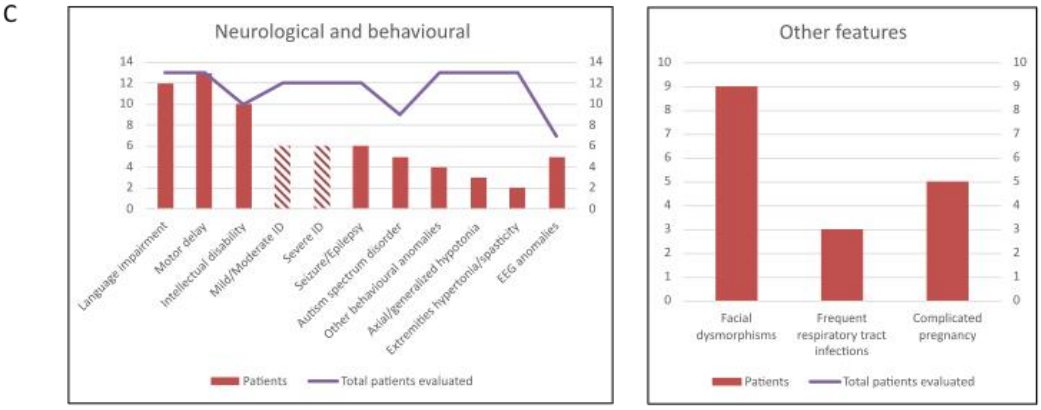
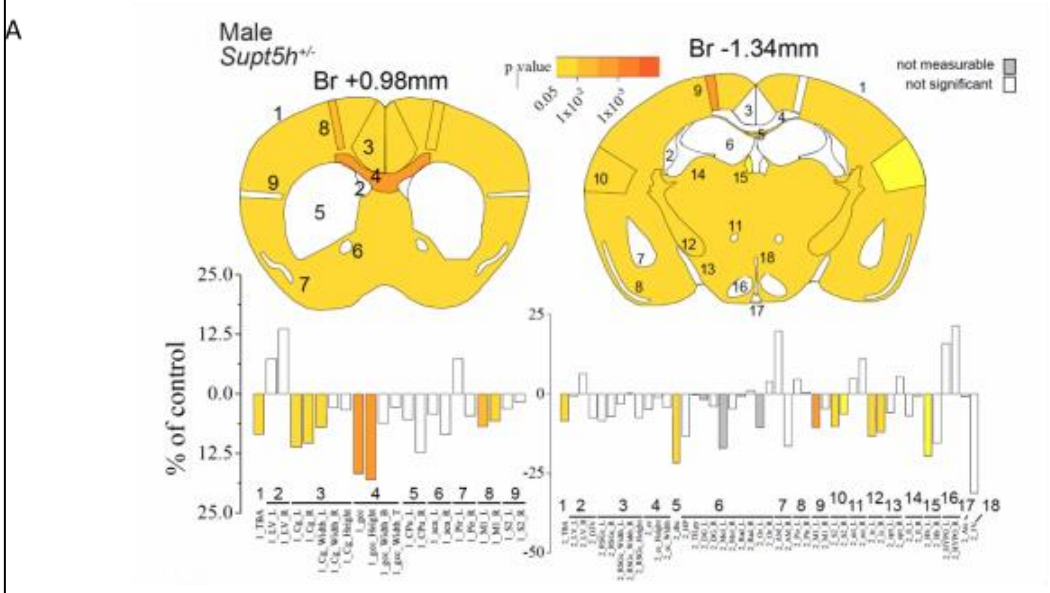
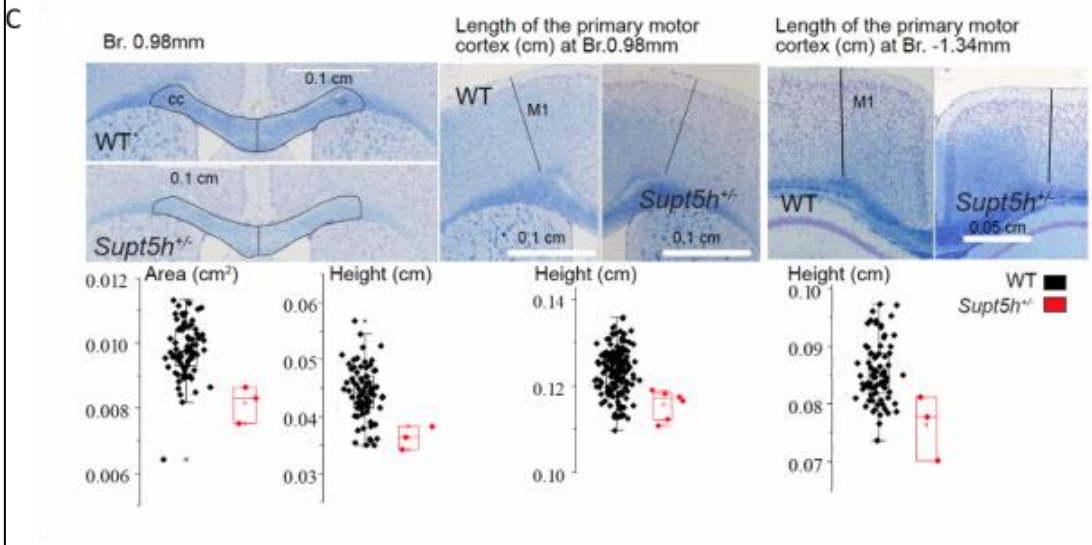

C
Neurological and behavioural
Language impairment
Motor delay
Intellectual disability
Mild/Moderate ID
Severe ID
Seizure/Epilepsy
Autism spectrum disorder
Other behavioural anomalies
Axial/generalized hypotonia
Extremities hypertonia/spasticity
EEG anomalies
Patients
Total patients evaluated
Other features
Facial dysmorphisms
Frequent respiratory tract infections
Complicated pregnancy
Patients
Total patients evaluated

Figure 2


A
Male
Supt5h+/-
Br +0.98mm
Br -1.34mm
p value
0.05
1x10-2
1x10-3
not measurable
not significant
% of control
25.0
12.5
0.0
12.5
25.0
25
0
-25
-50


C
Br. 0.98mm
Length of the primary motor cortex (cm) at Br.0.98mm
Length of the primary motor cortex (cm) at Br. -1.34mm
0.1 cm
cc
WT
Supt5h+/-
M1
0.05 cm
Area (cm²)
Height (cm)
WT
Supt5h+/-

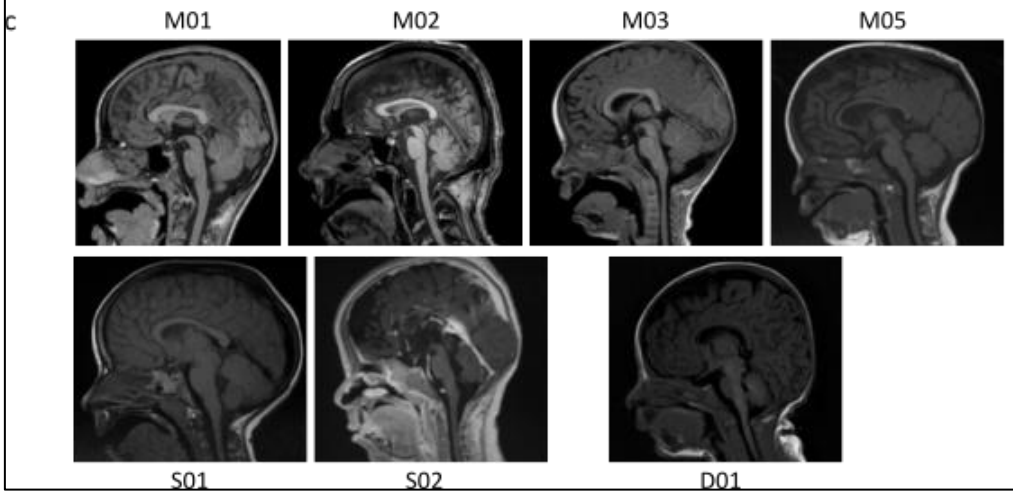
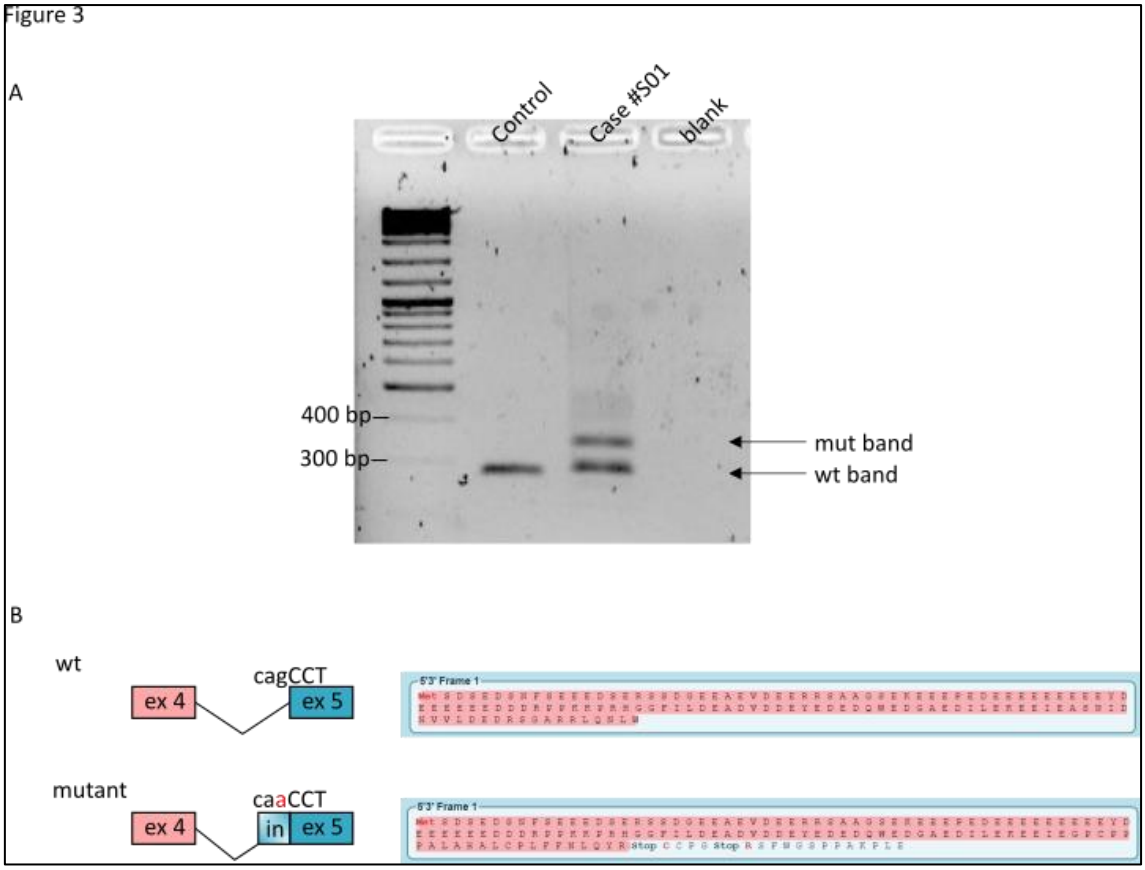

c
M01
M02
M03
M05
S01
S02
D01

Figure 3


A
Control
Case #S01
blank
400 bp
300 bp
mut band
wt band
B
wt
ex 4
cagCCT
ex 5
mutant
ex 4
caaCCT
in
ex 5

Conclusion and future perspectives

The NeuroWES project has shown to be a terrific "journey" for our group in the genetics of neurodevelopmental disorders. When the project started, the use of trio exome sequencing to find molecular diagnoses was still at its beginning, as well as the genetics of neurodevelopmental disorders. For these reasons, achieving a good diagnostic yield thanks to this analysis was already an important achievement.

However, the amount of data we obtained has been not only precious for a diagnostic purpose, but also for research ones.

Thanks to the NeuroWES project, we have been able to associate novel genes to different forms of NDDs, as in the case of *CAPRIN1* and *RPH3A*, described in Chapter 2. Describing a novel disease gene can be very intense and time demanding, and it can be problematic to collect detailed clinical information from patients in different geographic regions. Providing a functional characterization of the pathological mechanisms can be hard, particularly in cases where the genes are expressed only in certain tissues difficult to obtain (for instance the brain) or in the case in which the patients do not agree to donate their cells for *in vitro* studies. In both cases, *ad hoc* models could be necessary, and the use of complex cellular models - as hiPSCs - and/or genetic engineering with CRISPR/Cas could be necessary. However, the collection of a good number of patients with variants in a novel gene and functional studies aiming to clarify pathological mechanisms are fundamental for the progress of our knowledge.

The NeuroWES project though that every single patient should be analysed with great accuracy and attention. We found likely pathogenic or pathogenic variants in genes already associated with a disease but that did not explain the phenotype observed in our patients (Chapter 4). In those cases, it has been fundamental to collaborate with other research groups worldwide, to understand the role of variants in variable expressivity.

Solving the undiagnosed cases remain an important challenge. In our experience, the re-analysis of the undiagnosed cases after at least 12 months from the first analysis allowed to re-interpret missed pathogenic variants. The rationale is that the knowledge of NDD genetics is exponentially increasing, and every month novel disease genes are described: repeating the bioinformatics analysis could highlight novel described genes, lost months before. An example of this in our experience are the patients (family 1 and 2) described in Chapter 3, with variants in *TLK2* gene. If the re-analysis does not solve the case, other strategies (described in Chapter 1) can be used, and we are starting several projects for this purpose. One of the tasks for the next years will also be the identification of digenic and oligogenic disorders, which are expected to increase in the future.

A major issue in the diagnosis of NDDs is the relatively late onset of symptoms, representing a pitfall in the management of the patient and the choice of the correct therapy which is, however, often restricted to behavioural or educational intervention[266]. How can genetics contribute to the management of these disorders? First, genetic screening is important for recurrence risk assessment for the families. Then, the identification of a molecular diagnosis is necessary to foresee the onset of symptoms, applying early intervention strategies when possible and avoiding long and expensive procedures to find diagnosis[267] . Studies report that in up to 55% of cases genetic findings can lead to revised diagnosis and changes in clinical management[268,269]. This is particularly important for genetic NDDs for which pharmacological treatments are available and are subject of different response to treatment depending on the genetic molecular diagnosis, as in the case of some variants in *SCN2A* gene[270]. In the future, one of the main challenges will be to classify genetic causes into common pathways or mechanisms of action, to identify clusters of genes that can respond to the same treatments. For this purpose, many drug screenings and drug repositioning strategies are now object of study[271].

We are now in a "transition era" in which many changes are occurring in the field of genetics of NDDs and new challenges are arising: we are moving from the identification of new disease associated genes to their functional characterization and to the identification of potential treatments. However, while many efforts are being made in these directions, we also still need to keep on working on the improvement of pipelines of analysis and sequencing technologies. For instance, it has been largely shown how the use of long-read sequencing can increase (up to 20%) the yield of *de novo* mutations in patients already sequenced with short-read platforms[272–274]. Also, it has been reported how the combination of exome and genome sequencing with data repository analysis could enhance the identification of novel disease genes, suggesting a deep analysis of already available data can help to ameliorates our knowledge about the genetics of NDDs[275].

In conclusion, the genetics of NDDs is a rapidly evolving field that must deal with an enormous molecular and clinical heterogeneity. Linked to this, my most important take home message from my PhD (and from the NeuroWES study) is to never stop at the first negative result and keep on asking questions. Each case and each gene I encountered are unexplored worlds that need in-depth and accurate studying, and curiosity to be explored.

Appendix: supplementary data

Chapter 2: Novel genes involved in neurodevelopmental disorders

CAPRIN1, a novel gene involved in language impairment, ADHD and ASD

Supplementary can be found at https://doi.org/10.1093/brain/awac278*.*

Table of contents

RPH3A, a gene associated with a clinically variable neurodevelopmental disorder

Table of contents

Supplemental Materials and Methods

Whole exome sequencing, variant calling, and prioritization within the Autism Sequencing Consortium project

DNA was extracted from total blood using the ReliaPrep Blood gDNA Miniprep kit (Promega, Madi-son) following manufacturer's protocol and quantified with a NanoDrop spectrophotometer (Thermo Fisher Scientifics). Array-CGH was performed using a 60 K whole-genome oligonucleotide microarray (Agilent Technologies).

Patients were enrolled in the Autism Sequencing Consortium (ASC) (https://www.asc.broadinstitute.org/) project and their gDNA samples were sequenced at the Broad Institute on Illumina HiSeq sequencers as previously described[6,26].

ES raw data of the trio were processed and analyzed using an in-house implemented pipeline previously described[27,28] which is based on the GATK Best Practices[29]. The UCSC GRCh37/hg19 version of genome assembly was used as a reference for reads alignment by means of BWA-MEM[30] tool and the subsequent variant calling with HaplotypeCaller (GATK v3.8)[29]. We used

SnpEff v.5.0[31] and dbNSFP v.4.2[32] tools for variants functional annotation, including Combined Annotation Dependent Depletion (CADD) v.1.6[33], Mendelian Clinically Applicable Pathogenicity (M-CAP) v.1.0[200] and Intervar v.2.0.1 for functional impact prediction[34]. Thereby, the analysis was narrowed to variants which affect coding sequences or splice site regions. Moreover, high-quality variants were filtered against public databases (dbSNP150 and gnomAD ver.2.1.1) so that only variants with unknown frequency or having MAF <0.1%, as well as variants occurring with frequency < 1% in our population-matched database (~2800 exomes) were considered.

A further variant stratification in conformity with the American College of Medical Genetics and Genomics (ACMG) guideline[34], considering also mode of inheritance and functional in silico prediction of impact, allowed us considering the final set of variants for possible associations with the phenotype. All variants are referred to GRCh37 annotation and to NM_ 001143854.2. Identified variants were confirmed by Sanger sequencing using standard conditions and/or the Integrative Genomics Viewer browser (IGV)[276].

Variants were analyzed with the VarSome tool (https://www.varsome.com/)[159] as a starting point for further analysis. This allowed evaluation of at least nine *in silico* predictors simultaneously. Variants frequencies were evaluated using Genome Aggregation Database (GnomAD, ver 2.1.1).

Alignments between species were performed using CLUSTALW Multiple Sequence Alignment tool (https://www.genome.jp/tools-bin/clustalw), with standard setting and slow pairwise alignment[277].

Cell culture

HEK293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FBS, 1% Pen-Strep and 1 mM sodium pyruvate (Thermo Fisher Scientific, Waltham, MA, USA) at 37°C, 5% CO_2.

Minigene assay generation and expression

Minigene constructs were synthesized by VectorBuilder (Guangzhou, China), using the pSPL3 exon trapping vector (Invitrogen, Thermo Fisher Scientific). Plasmids have been deposited on Addgene with the codes 190476, 190477, 190478 and 190479. HEK293 cells were plated on a 12-wells plate at a confluence of $3.5x10^5$ cells/well. The day after, with a confluence of about 70%, cells were transfected with pSPL3 constructs (1 µg/construct) using Lipofectamine 2000 (Thermo Fisher Scientific) and Opti-MEM Reduced Serum medium. After 48 hours, cells were detached using trypsin-EDTA (0.25%) (Gibco), and TRIFast (EuroClone, Pero, MI, Italy) was added to the cell pellet.

Total RNA was extracted using the Direct-Zol RNA MiniPrep system (Zymo Research, Irvine, CA,

USA); genomic DNA was removed by treatment with DNase I (Sigma-Aldrich), following the manufacturer's protocol. Complementary DNA (cDNA) was generated using the M-MLV Reverse Transcriptase kit (Invitrogen, Thermo Fisher Scientific). cDNA was amplified with touchdown PCR using KAPA Taq PCR Kit (Roche Diagnostics, Basel Switzerland). The following primers were used: 5'-TCTGAGTCACCTGGACAACC and 5'- ATCTCAGTGGTATTTGTGAGC. Amplimers were visualized into 2% TBE-agarose gel using ChemiDoc Imaging System (Bio Rad, Hercules, CA, USA) and sequenced by Sanger method after phosphates removal using FastAP Thermosensitive Alkaline Phosphatase (Thermo Fisher Scientific).

Cell cultures, transfection, plasmids

Hippocampal primary neuronal cultures were prepared from embryonic day 18-19 (E18-E19) Sprague Dawley rat hippocampi (Charles Rivers, Milan, Italy) as previously described[278]. Neurons were transfected on day in vitro 8 (DIV8) through the calcium-phosphate with 2-5 µg of plasmid DNA. The GFP plasmid was kindly provided by Dr M. Passafaro (Università di Milano – Bicocca, Italy). The hSyn-RFP-Rph3A-WPRE plasmids were purchased from Addgene, the hSyn-RFP-Rph3AN605S-WPRE and hSyn-RFP-Rph3AN605S-WPRE were produced and purchased from Vector Builder. Neurons were used for the various experiment on DIV16.

Antibodies

For spine morphology experiment, the following primary antibodies were used: mouse anti-RFP (OAEA00012, Aviva; dilution 1:500 ICC) and chicken anti-GFP (AB16901, Millipore; dilution 1:300 ICC). The following secondary antibodies were used: goat anti-mouse-AlexaFluor555 (A-21424, Invitrogen; dilution 1:1000 ICC) and goat anti-chicken-AlexaFluor488 (13C0523, Immunological sciences; dilution 1:1000 ICC).

For Surface/Total assay, the following primary antibodies were used: rabbit anti-N-term-GluN2A (480031, Invitrogen; dilution 1:100 ICC), mouse anti-GluA1 (75-327, Neuromab; dilution 1:100 ICC) and mouse anti-RFP (OAEA00012, Aviva; dilution 1:500 ICC). The following secondary antibodies were used: goat anti-rabbit-AlexaFluor488 (A21206, Life Technologies; dilution 1:1000 ICC), goat anti-rabbit-AlexaFluor647 (A21245, Life Technologies; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor555 (A-21424, Invitrogen; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor488 (411029, Invitrogen; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor647 (A21235, Invitrogen; dilution 1:1000).

For colocalization in Airyscan modality experiment, the following primary antibodies were used: mouse anti-PSD-95 (192757, Abcam; dilution 1:1000 ICC) and rabbit anti-N-term-GluN2A (480031, Invitrogen; dilution 1:100 ICC). The following secondary antibodies were used: goat

anti-mouse-AlexaFluor647 (A21235, Invitrogen; dilution 1:1000) and goat anti-rabbit-AlexaFluor488 (A21206, Life Technologies; dilution 1:1000 ICC).

Computational procedures

All the computational procedures were carried out through the Molecular Operating Environment (MOE) version 2020.09 and the Amber10:EHT forcefield. To build a three-dimensional model of human Rph3A C2A domain, the crystallographic structure of the C2A domain of Rph3A of Rattus norvegicus was selected as template (PDB ID: 4LT7). Since this structure is co-crystallized with only one Ca2+ ion, we imported the two Ca2+ ions from Rph3A of Mus musculus (PDB ID:2k3h), during homology modeling procedures, to reshape the calcium binding pocket accordingly. To build the three-dimensional model of human Rph3A C2B domain, we selected the NMR structure of the C2B domain of Rph3A of Rattus norvegicus as template (PDB ID: 2CM6).

Rph3AT450S and Rph3AN618S mutants were generated using the MOE Protein Design module; for each mutation, an ensemble of protein conformations was generated using the Low Mode molecular dynamics (MD) approach. The change in stability due to the introduction of mutations is computed as the difference in free energy of folding of the mutant and the wild type forms, using the FOLDEF (for FOLD-X energy function) algorithm, as implemented in MOE[279]. The MOE Protein Builder module was used to study the Rph3AT450S rotamers from an energy point of view.

Data Presentation and Statistical Analysis

Confocal Imaging: All the group values are expressed as mean ± SEM. Normality was examined using D'Agostino-Pearson or Shapiro-Wilk tests. Comparisons between groups were performed using the following tests as appropriate: two-tailed unpaired Student's t-test, Mann-Whitney U test, one-way analysis of variance (ANOVA) followed by Dunnett *post-hoc* tests or Kruskal-Wallis' test followed by Dunn's *post-hoc* test. Outliers were identified with the ROUT method (Q = 5%). Significance was defined as $p < 0.05$. All statistical analyses were done using the GraphPad Prism statistical package (GraphPad software). Sample sizes for the specific types of experiments conducted are like those generally employed in the field. When appropriate, experiments were performed under blind conditions.

Electrophysiology: Data are given as mean ± SEM for number of cells. The normal distribution of experimental data was assessed through the Shapiro-Wilk's normality test. The statistical significance of our data has been evaluated by considering two sample groups normally distributed and applying the two-way ANOVA test followed by Bonferroni's test.

Antibodies

For spine morphology experiment, the following primary antibodies were used: mouse anti-RFP (OAEA00012, Aviva; dilution 1:500 ICC) and chicken anti-GFP (AB16901, Millipore; dilution 1:300 ICC). The following secondary antibodies were used: goat anti-mouse-AlexaFluor555 (A-21424, Invitrogen; dilution 1:1000 ICC) and goat anti-chicken-AlexaFluor488 (13C0523, Immunological sciences; dilution 1:1000 ICC).

For Surface/Total assay, the following primary antibodies were used: rabbit anti-N-term-GluN2A (480031, Invitrogen; dilution 1:100 ICC), mouse anti-GluA1 (75-327, Neuromab; dilution 1:100 ICC) and mouse anti-RFP (OAEA00012, Aviva; dilution 1:500 ICC). The following secondary antibodies were used: goat anti-rabbit-AlexaFluor488 (A21206, Life Technologies; dilution 1:1000 ICC), goat anti-rabbit-AlexaFluor647 (A21245, Life Technologies; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor555 (A-21424, Invitrogen; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor488 (411029, Invitrogen; dilution 1:1000 ICC), goat anti-mouse-AlexaFluor647 (A21235, Invitrogen; dilution 1:1000).

For colocalization in Airyscan modality experiment, the following primary antibodies were used: mouse anti-PSD-95 (192757, Abcam; dilution 1:1000 ICC) and rabbit anti-N-term-GluN2A (480031, Invitrogen; dilution 1:100 ICC). The following secondary antibodies were used: goat anti-mouse-AlexaFluor647 (A21235, Invitrogen; dilution 1:1000) and goat anti-rabbit-AlexaFluor488 (A21206, Life Technologies; dilution 1:1000 ICC).

Data Presentation and Statistical Analysis

Confocal Imaging: All the group values are expressed as mean ± SEM. Normality was examined using D'Agostino-Pearson or Shapiro-Wilk tests. Comparisons between groups were performed using the following tests as appropriate: two-tailed unpaired Student's t-test, Mann-Whitney U test, one-way analysis of variance (ANOVA) followed by Dunnett *post-hoc* tests or Kruskal-Wallis' test followed by Dunn's *post-hoc* test. Outliers were identified with the ROUT method (Q = 5%). Significance was defined as $p < 0.05$. All statistical analyses were done using the GraphPad Prism statistical package (GraphPad software). Sample sizes for the specific types of experiments conducted are like those generally employed in the field. When appropriate, experiments were performed under blind conditions.

Electrophysiology: Data are given as mean ± SEM for number of cells. The normal distribution of experimental data was assessed through the Shapiro-Wilk's normality test. The statistical significance of our data has been evaluated by considering two sample groups normally distributed and applying the two-way ANOVA test followed by Bonferroni's test.

Cases report

Patient #1

Patient #1 was a 12-years old White male born at full term. At 4 years of age, he showed moderate intellectual disability. At last clinical evaluation at 11 years, his weight was 62 kg (99th percentile, +2.5 SD), height 153 cm (91st percentile, +.133 SD), and occipitofrontal circumference 58.5 cm (>99th percentile, +3.7 SD). He was following the middle school without any special need, and he showed high functioning autism spectrum disorder, dysgraphia, and hyperactivity. He showed mild facial dysmorphisms: low anterior hairline, hypertelorism, down-slanted palpebral fissures, anteverted nares, low hanging columella, short philtrum, and widely spaced incisors. Additionally, inverted nipples, back hair, tapering fingers, recurrent upper respiratory tract infections, nasal speech and hyperphagia were reported.

Trio-based whole exome sequencing identified the *de novo* variant c.1853A>G; p.(Asn618Ser) in the *RPH3A* gene.

Patient #2

Patient #2 was a white Hispanic/American Indian 5-years old male, born full term by Caesarean section. The birth was complicated by feeding problems and incoordination, gastroesophageal reflux (GERD), noisy breathing/stridor, hypotonia, failure to thrive, diarrhoea, questionable seizure and need for G-tube placement. His birth weight was 3.47 kg (44th percentile, -0.15 SD), length 51 cm (62nd percentile, +0.32 SD), occipitofrontal circumference 36 cm (52nd percentile, +0.06 SD). At the last examination (5 years old), his weight was 18.6 kg (53rd percentile, +0.07 SD) and height 104 cm (15th percentile, -1.04 SD). He was wheelchair dependent, with general hypotonia and limb hypertonia. Speech was impaired, with inability to speak; moderate intellectual disability and seizures were also reported. Brain MRI showed periventricular leukomalacia with no evidence of haemorrhage. Breathing problems, including wheezing, chronic night cough, rattly breathing and obstructive sleep apnoea were reported. Additionally, he showed ocular problems (discharge, photophobia, strabismus), gastrointestinal issues (poor sucking and swallowing, G-tube, GERD), sleeping disorder (night terrors, snoring) and genito-urinary issues (hypospadias). Facial dysmorphisms were reported, specifically prominent forehead, sparse eyebrows, almond-shaped palpebral fissure, epicanthus, down-slanted palpebral fissures, depressed nasal bridge, wide nasal bridge, microtia (ear length 3cm), low-set ears bilaterally, smooth philtrum, thin upper lip, downturned corners of mouth, pointed chin. Pictures were not available.

Genetic testing by whole genome sequencing identified the *de novo* variant *RPH3A* c.1349C>G p.(Thr450Ser), and the following variants which may be related to the phenotype:

- maternally-inherited VUS X-linked hemizygous variant in *CUL4B* gene [NM_001079872.2, c.2523G>C; p.(Lys841Asn)]; this gene is associated with "Cabezas syndrome, characterized primarily by short stature, hypogonadism, and abnormal gait, with other more variable features such as speech delay, prominent lower lip, and tremor (MIM# 300354); the variant is shared with his brother who has a history of developmental delay, malnutrition, and wheelchair dependence.
- maternally-inherited VUS heterozygous variant in *PRKAG2* [NM_016203.4, c.547G>A; p.(Glu183Lys]; proband was indeed diagnosed with Wolff-Parkinson-White syndrome (MIM# 194200), a form of hypertrophic cardiomyopathy.
- maternally-inherited pathogenic heterozygous variant in *SCN4A* (NM_000334.4, c.3938C>T; p.Thr1313Met; ClinVar 20943), causing paramyotonia congenita (MIM# 168300).

Variants in *CUL4B*, *PRKAG2* and *SCN4A* gene can be associated with the motor impairment, language delay and hypo-hypertonia, but they cannot explain the seizures.

Patient #3

Patient #3 was a 37-years old White male born full term by Caesarean section. He showed low average performance, with pervasive disturbance maximally implicating dominant fronto-temporal regions. At 13 years of age, he started to show drug-resistant focal to bilateral tonic-clonic seizures. His EEG showed bitemporal independent epileptiform discharges and independent bitemporal seizure onset. A brain MRI showed extensive left temporal focal cortical dysplasia a left hippocampal sclerosis. The variant *RPH3A* c.626G>A; p.(Arg209Lys) was identified by whole genome sequencing within the 100,000 Genomes Project[20].

Patient #4

Clinical information for case #4 were very limited, but he was reported to show childhood autism spectrum disorder, intellectual disability, obesity and inverted nipples. His genetic testing, performed by whole genome sequencing within the 100,000 Genomes Project[20], reported the missense variant in *RPH3A* c.705G>T; p.(Arg235Ser)

Patient #5

Patient #5 was a 59-years old White male born at full term. He had moderate intellectual disability, attended "special needs" school, and had impaired verbal abilities. At 1 year old, he started to show focal to bilateral tonic-clonic seizures, which did not respond to any drug

treatment. EEG exam showed bilateral independent anterior temporal sharp waves. MRI reported bilateral hippocampal abnormalities, namely right hippocampal sclerosis, and left bulky and abnormal signal. His genetic testing, performed by whole genome sequencing within the 100,000 Genomes Project[20], reported the missense variant in *RPH3A* c.1524G>C; p.(Gln508His).

Table S1. Bioinformatics predictions of *RPH3A* variants

Patient	cDNA	Protein	Inheritance	CADD 1.6	Aggregated prediction (Frank	REVEL	MetaLR	Mut.Taste	Mut Assessor	FATHMM	SIFT	GERP	GenoCanyon	fitCons	SNP&Go
Case 1	c.1853A>G	p.(Asn618Ser)	*de novo*	26,4	Deleterious	Deleterio	Uncertain	Disease	Medium	Uncertain	Uncertain	5,42	Deleterious	Benign	Neutral
Case 2	c.1349C>G	p.(Thr450Ser)	*de novo*	26,6	Deleterious	Deleterio	Uncertain	Disease	Medium	Uncertain	Uncertain	5,43	Deleterious	Deleterious	Disease
Case 3	c.626G>A	p.(Arg209Lys)	n.d.	17	Uncertain	Benign	Benign	Disease	Neutral	Uncertain	Benign	3,96	Benign	Benign	Neutral
Case 4	c.705G>T	p.(Arg235Ser)	not maternal	19,5	Uncertain	Benign	Benign	Benign	Medium	Uncertain	Benign	4,82	Benign	Benign	Neutral
Case 5	c.1524G>C	p.(Gln508His)	n.d.	22	Benign	Benign	Benign	Disease	Medium	Uncertain	Uncertain	4,31	Deleterious	Benign	Neutral
Note: (NM_001143854.2)															

Legends to supplementary figures

Figure S1. Minigene assay of variant c.444G>T. **(A)** Localization of the variant predicted to cause the Glu148Asp change. **(B)** Agarose gel analysis of cDNA, showing a higher band for 444T construct compared to 444G. **(C)** Sanger sequencing of 444G and 444T constructs, showing a partial retaining of intron 7 with 444T. The retained sequence was GTGAGTGCCCTGGTCCCACCTGGTGCCTAGATCACCCTCCTTTCTTGGCCAGCTTAAGAGGTGCCTTAA GAG, which is predicted to generate three different stop codons.

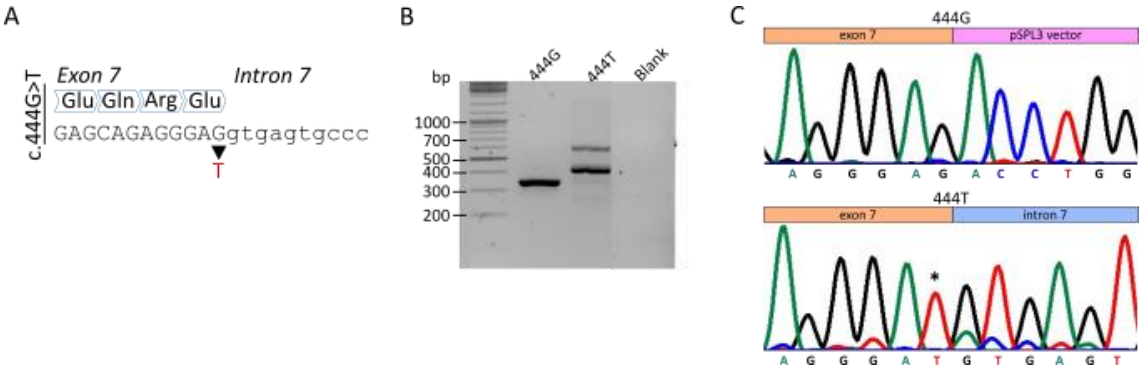
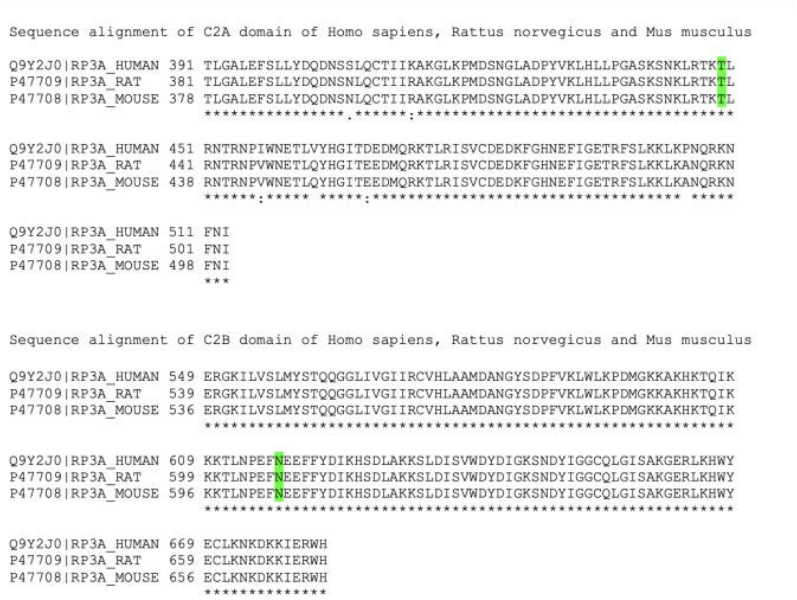


Figure S2. Minigene assay of variant c.1853A>G. **(A)** Localization of the variant causing p.(Asn618Ser) change. **(B)** Agarose gel analysis of cDNA, showing no differences were visible for 1853A construct compared to 1853G.

Figure S3. Sequence alignment of *in vitro* analysed variants. Sequence alignment among the Rph3A C2A and C2B domains of *Homo sapiens*, *Rattus Norvegicus* and *Mus musculus*. Rph3A^{T450S} and Rph3A^{N618S} mutations are highlighted in green. Amino acid Thr450 of the human protein corresponds to Thr437 in mouse and Thr440 of rat proteins, while Asn618 of the human form corresponds respectively to Asn605 of the mouse and Asn607 of the rat proteins.



Chapter 3: New mechanisms for known genes

TLK2-related neurodevelopmental disorder: new insights on the pathological mechanisms

Supplementary can be found at http://dx.doi.org/10.1136/jmedgenet-2020-107281.

Table of contents

Chapter 4: Novel phenotypes associated to known genes

KCNK18: biallelic variants altering TRESK channel activity are associated with intellectual disability and neurodevelopmental disorders

Supplementary can be found at https://doi.org/10.3390/ijms22116064.

Table of contents

Supplementary table

WDR26 variants: expanding the clinical phenotype in two additional cases

Supplementary can be found at https://doi.org/10.1002/ajmg.a.62157.

Table of contents

References

1. Gilissen C, Hehir-Kwa JY, Thung DT, et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature*. 2014;511(7509):344-347. doi:10.1038/nature13394

2. Parenti I, Rabaneda LG, Schoen H, Novarino G. Neurodevelopmental Disorders: From Genetics to Functional Pathways. *Trends Neurosci*. 2020;43(8):608-621. doi:10.1016/j.tins.2020.05.004

3. Savatt JM, Myers SM. Genetic Testing in Neurodevelopmental Disorders. *Front Pediatr*. 2021;9. doi:10.3389/fped.2021.526779

4. Blesson A, Cohen JS. Genetic Counseling in Neurodevelopmental Disorders. *Cold Spring Harb Perspect Med*. 2020;10(4):a036533. doi:10.1101/cshperspect.a036533

5. Fu JM, Satterstrom FK, Peng M, et al. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat Genet*. 2022;54(9):1320-1331. doi:10.1038/s41588-022-01104-0

6. de Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. Published online 2014. doi:10.1038/nature13772

7. de Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 2014;515(7526):209-215. doi:10.1038/nature13772

8. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

9. Collins RL, Glessner JT, Porcu E, et al. A cross-disorder dosage sensitivity map of the human genome. *Cell*. 2022;185(16):3041-3055.e25. doi:10.1016/j.cell.2022.06.036

10. Iossifov I, Levy D, Allen J, et al. Low load for disruptive mutations in autism genes and their biased transmission. *Proceedings of the National Academy of Sciences*. 2015;112(41). doi:10.1073/pnas.1516376112

11. Ruzzo EK, Pérez-Cano L, Jung JY, et al. Inherited and De Novo Genetic Risk for Autism Impacts Shared Networks. *Cell*. 2019;178(4):850-866.e26. doi:10.1016/j.cell.2019.07.015

12. Anney R, Klei L, Pinto D, et al. Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum Mol Genet*. Published online 2012. doi:10.1093/hmg/dds301

13. Niemi MEK, Martin HC, Rice DL, et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*. 2018;562(7726):268-271. doi:10.1038/s41586-018-0566-4

14. Pizzo L, Jensen M, Polyak A, et al. Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants. *Genet Med*. 2019;21(4):816-825. doi:10.1038/s41436-018-0266-3

15. Dudbridge F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genet*. 2013;9(3):e1003348. doi:10.1371/journal.pgen.1003348

16. Lord C, Brugha TS, Charman T, et al. Autism spectrum disorder. *Nat Rev Dis Primers*. 2020;6(1):5. doi:10.1038/s41572-019-0138-4

17. Tran NQV, Miyake K. Neurodevelopmental Disorders and Environmental Toxicants: Epigenetics as an Underlying Mechanism. *Int J Genomics*. 2017;2017:7526592. doi:10.1155/2017/7526592

18. Lintas C. Linking genetics to epigenetics: The role of folate and folate-related pathways in neurodevelopmental disorders. *Clin Genet*. 2019;95(2):241-252. doi:10.1111/cge.13421

19. Warrier V, Zhang X, Reed P, et al. Genetic correlates of phenotypic heterogeneity in autism. *Nat Genet*. 2022;54(9):1293-1304. doi:10.1038/s41588-022-01072-5

20. Berecki G, Howell KB, Heighway J, et al. Functional correlates of clinical phenotype and severity in recurrent SCN2A variants. *Commun Biol*. 2022;5(1):515. doi:10.1038/s42003-022-03454-1

21. Pavinato L, Delle Vedove A, Carli D, et al. CAPRIN1 haploinsufficiency causes a neurodevelopmental disorder with language impairment, ADHD and ASD. *Brain*. Published online July 27, 2022. doi:10.1093/brain/awac278

22. Reijnders MRF, Miller KA, Alvi M, et al. 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;102(6):1195-1203. doi:10.1016/j.ajhg.2018.04.014

23. Pavinato L, Villamor-Payà M, Sanchiz-Calvo M, et al. Functional analysis of TLK2 variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis. *J Med Genet*. Published online 2020. doi:10.1136/jmedgenet-2020-107281

24. Buxbaum JD, Daly MJ, Devlin B, Lehner T, Roeder K, State MW. The Autism Sequencing Consortium: Large-Scale, High-Throughput Sequencing in Autism Spectrum Disorders. *Neuron*. 2012;76(6):1052-1056. doi:10.1016/j.neuron.2012.12.008

25. Köhler S, Gargano M, Matentzoglu N, et al. The Human Phenotype Ontology in 2021. *Nucleic Acids Res*. 2021;49(D1):D1207-D1217. doi:10.1093/nar/gkaa1043

26. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. Published online 2020. doi:10.1016/j.cell.2019.12.036

27. Bauer CK, Calligari P, Radio FC, et al. Mutations in KCNK4 that Affect Gating Cause a Recognizable Neurodevelopmental Syndrome. *Am J Hum Genet*. Published online 2018. doi:10.1016/j.ajhg.2018.09.001

28. Flex E, Martinelli S, van Dijck A, et al. Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. *Am J Hum Genet*. Published online 2019. doi:10.1016/j.ajhg.2019.07.007

29. van der Auwera GA, Carneiro MO, Hartl C, et al. From fastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. Published online 2013. doi:10.1002/0471250953.bi1110s43

30. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. Published online 2009. doi:10.1093/bioinformatics/btp324

31. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. Published online 2012. doi:10.4161/fly.19695

32. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat*. Published online 2016. doi:10.1002/humu.22932

33. Kircher M, Witten DM, Jain P, O'roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. Published online 2014. doi:10.1038/ng.2892

34. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet*. Published online 2017. doi:10.1016/j.ajhg.2017.01.004

35. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*. Published online 2015. doi:10.1038/gim.2015.30

36. Pavinato L, Trajkova S, Grosso E, et al. Expanding the clinical phenotype of the ultra-rare Skraban-Deardorffsyndrome: Two novel individuals with WDR26 loss-of-function variants and a literature review. *Am J Med Genet A*. 2021;185(6):1712-1720. doi:10.1002/ajmg.a.62157

37. Verdura E, Rodríguez-Palmero A, Vélez-Santamaria V, et al. Biallelic PI4KA variants cause a novel neurodevelopmental syndrome with hypomyelinating leukodystrophy. *Brain*. 2021;144(9):2659-2669. doi:10.1093/brain/awab124

38. Aref-Eshghi E, Kerkhof J, Pedro VP, et al. Evaluation of DNA Methylation Episignatures for Diagnosis and Phenotype Correlations in 42 Mendelian Neurodevelopmental Disorders. *The American Journal of Human Genetics*. 2021;108(6):1161-1163. doi:10.1016/j.ajhg.2021.04.022

39. Levy MA, McConkey H, Kerkhof J, et al. Novel diagnostic DNA methylation episignatures expand and refine the epigenetic landscapes of Mendelian disorders. *Human Genetics and Genomics Advances*. 2022;3(1):100075. doi:10.1016/j.xhgg.2021.100075

40. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. *Hum Mutat*. Published online 2015. doi:10.1002/humu.22844

41. Caulfield M, DDavies J, Dennys M, Elbahy L, Fowler T, Hill S. *The National Genomics Research and Healthcare Knowledgebase Amendment to the 100,000 Genomes Project Protocol V4*.; 2017. doi:https://doi.org/10.6084/m9.figshare.4530893.v5

42. Marcogliese PC, Deal SL, Andrews J, et al. Drosophila functional screening of de novo variants in autism uncovers damaging variants and facilitates discovery of rare neurodevelopmental diseases. *Cell Rep*. 2022;38(11). doi:10.1016/j.celrep.2022.110517

43. Damianidou E, Mouratidou L, Kyrousi C. Research models of neurodevelopmental disorders: The right model in the right place. *Front Neurosci*. 2022;16. doi:10.3389/fnins.2022.1031075

44. Zhao X, Bhattacharyya A. Human Models Are Needed for Studying Human Neurodevelopmental Disorders. *The American Journal of Human Genetics*. 2018;103(6):829-857. doi:10.1016/j.ajhg.2018.10.009

45. Elitt MS, Barbar L, Tesar PJ. Drug screening for human genetic diseases using iPSC models. *Hum Mol Genet*. 2018;27(2):R89-R98. doi:10.1093/hmg/ddy186

46. Qian X, Song H, Ming G li. Brain organoids: advances, applications and challenges. *Development*. 2019;146(8). doi:10.1242/dev.166074

47. Watanabe M, Buth JE, Vishlaghi N, et al. Self-Organized Cerebral Organoids with Human-Specific Features Predict Effective Drugs to Combat Zika Virus Infection. *Cell Rep*. 2017;21(2):517-532. doi:10.1016/j.celrep.2017.09.047

48. Liu P, Meng L, Normand EA, et al. Reanalysis of Clinical Exome Sequencing Data. *New England Journal of Medicine*. 2019;380(25):2478-2480. doi:10.1056/NEJMc1812033

49. Ewans LJ, Schofield D, Shrestha R, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genetics in Medicine*. 2018;20(12):1564-1574. doi:10.1038/gim.2018.39

50. Costain G, Jobling R, Walker S, et al. Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. *European Journal of Human Genetics*. 2018;26(5):740-744. doi:10.1038/s41431-018-0114-6

51. Giovenino C, Trajkova S, Pavinato L, et al. Skewed X-chromosome Inactivation in Unsolved Neurodevelopmental Disease Cases Can Guide Re-evaluation for X-linked Genes. doi:10.21203/rs.3.rs-2179710/v1

52. Chaves LD, Carvalho LML, Tolezano GC, et al. Skewed X-chromosome inactivation in women with idiopathic intellectual disability as indicative of pathogenic variants. doi:10.21203/rs.3.rs-1859222/v1

53. Friederich MW, Timal S, Powell CA, et al. Pathogenic variants in glutamyl-tRNAGln amidotransferase subunits cause a lethal mitochondrial cardiomyopathy disorder. *Nat Commun*. 2018;9(1):4065. doi:10.1038/s41467-018-06250-w

54. Kamps R, Szklarczyk R, Theunissen TE, et al. Genetic defects in mtDNA-encoded protein translation cause pediatric, mitochondrial cardiomyopathy with early-onset brain disease. *Eur J Hum Genet*. 2018;26(4):537-551. doi:10.1038/s41431-017-0058-2

55. Kohda M, Tokuzawa Y, Kishita Y, et al. A Comprehensive Genomic Analysis Reveals the Genetic Landscape of Mitochondrial Respiratory Chain Complex Deficiencies. *PLoS Genet*. 2016;12(1):e1005679. doi:10.1371/journal.pgen.1005679

56. Dursun F, Genc HM, Mine Yılmaz A, et al. Primary adrenal insufficiency in a patient with biallelic QRSL1 mutations. *Eur J Endocrinol*. 2022;187(3):K27-K32. doi:10.1530/EJE-22-0233

57. Yang ZS, Qing H, Gui H, Luo J, Dai LJ, Wang B. Role of caprin-1 in carcinogenesis (Review). *Oncol Lett*. 2019;18(1):15-21. doi:10.3892/ol.2019.10295

58. Shiina N, Yamaguchi K, Tokunaga M. RNG105 deficiency impairs the dendritic localization of mRNAs for Na +/K+ ATPase subunit isoforms and leads to the degeneration of neuronal networks. *Journal of Neuroscience*. Published online 2010. doi:10.1523/JNEUROSCI.6386-09.2010

59. Ohashi R, Takao K, Miyakawa T, Shiina N. Comprehensive behavioral analysis of RNG105 (Caprin1) heterozygous mice: Reduced social interaction and attenuated response to novelty. *Sci Rep*. Published online 2016. doi:10.1038/srep20775

60. Nakayama K, Ohashi R, Shinoda Y, et al. RNG105/caprin1, an RNA granule protein for dendritic mRNA localization, is essential for long-term memory formation. *Elife*. 2017;6:1-32. doi:10.7554/eLife.29677

61. el Fatimy R, Tremblay S, Dury AY, et al. Fragile mental retardation protein interacts with the RNA-binding protein Caprin1 in neuronal

ribonucleoprotein complexes. *PLoS One*. 2012;7(6). doi:10.1371/journal.pone.0039338

62. Wu Y, Zhu J, Huang X, Du Z. Crystal structure of a dimerization domain of human Caprin-1: Insights into the assembly of an evolutionarily conserved ribonucleoprotein complex consisting of Caprin-1, FMRP and G3BP1: Insights. *Acta Crystallogr D Struct Biol*. 2016;72(6):718-727. doi:10.1107/S2059798316004903

63. Solomon S, Xu Y, Wang B, et al. Distinct Structural Features ofCaprin-1 Mediate Its Interaction with G3BP-1 and Its Induction of Phosphorylation of Eukaryotic Translation InitiationFactor 2α, Entry to Cytoplasmic Stress Granules, and Selective Interaction with a Subset of mRNAs. *Mol Cell Biol*. 2007;27(6):2324-2342. doi:10.1128/MCB.02300-06

64. Shiina N, Shinkura K, Tokunaga M. A novel RNA-binding protein in neuronal RNA granules: Regulatory machinery for local translation. *Journal of Neuroscience*. Published online 2005. doi:10.1523/JNEUROSCI.0382-05.2005

65. Solomon S, Xu Y, Wang B, et al. Distinct Structural Features ofCaprin-1 Mediate Its Interaction with G3BP-1 and Its Induction of Phosphorylation of Eukaryotic Translation Initiation Factor 2α, Entry to Cytoplasmic Stress Granules, and Selective Interaction with a Subset of mRNAs. *Mol Cell Biol*. 2007;27(6):2324-2342. doi:10.1128/mcb.02300-06

66. Deciphering Developmental Disorders S, McRae JF, Clayton S, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433-438. doi:10.1038/nature21062

67. Fu JM, Satterstrom FK, Peng M, et al. Rare coding variation illuminates the allelic architecture, risk genes, cellular expression patterns, and phenotypic context of autism. *medRxiv*. Published online January 1, 2021:2021.12.20.21267194. doi:10.1101/2021.12.20.21267194

68. Matz P, Adjaye J. Generation of iPSC line epiHUVEC from human umbilical vein endothelial cells. *Stem Cell Res*. 2015;15(3):581-583. doi:10.1016/j.scr.2015.10.004

69. Frank S, Ahuja G, Bartsch D, et al. yylncT Defines a Class of Divergently Transcribed lncRNAs and Safeguards the T-mediated Mesodermal Commitment of Human PSCs. *Cell Stem Cell*. 2019;24(2):318-327.e8. doi:10.1016/j.stem.2018.11.005

70. Boissart C, Poulet A, Georges P, et al. Differentiation from human pluripotent stem cells of cortical neurons of the superficial layers amenable to psychiatric disease modeling and high-throughput drug screening. *Transl Psychiatry*. 2013;3. doi:10.1038/tp.2013.71

71. Darville H, Poulet A, Rodet-Amsellem F, et al. Human Pluripotent Stem Cell-derived Cortical Neurons for High Throughput Medication Screening in Autism: A Proof of Concept Study in SHANK3 Haploinsufficiency Syndrome. *EBioMedicine*. 2016;9:293-305. doi:10.1016/j.ebiom.2016.05.032

72. Bachmann S, Linde J, Bell M, Spehr M, Zempel H, Zimmer-Bensch G. DNA Methyltransferase 1 (DNMT1) Shapes Neuronal Activity of Human iPSC-Derived Glutamatergic Cortical Neurons. *Int J Mol Sci*. 2021;22(4):2034. doi:10.3390/ijms22042034

73. Schmidt EK, Clavarino G, Ceppi M, Pierre P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods*. 2009;6(4):275-277. doi:10.1038/nmeth.1314

74. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*. Published online 2015. doi:10.1038/nprot.2015.053

75. Reynolds CR, Islam SA, Sternberg MJE. EzMol: A Web Server Wizard for the Rapid Visualization and Image Production of Protein and Nucleic Acid Structures. *J Mol Biol*. Published online 2018. doi:10.1016/j.jmb.2018.01.013

76. Kozakov D, Beglov D, Bohnuud T, et al. How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics*. 2013;81(12):2159-2166. doi:10.1002/prot.24403

77. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein–protein docking. *Nat Protoc*. 2017;12(2):255-278. doi:10.1038/nprot.2016.169

78. Vajda S, Yueh C, Beglov D, et al. New additions to the <scp>C</scp> lus <scp>P</scp> ro server motivated by <scp>CAPRI</scp>. *Proteins: Structure, Function, and Bioinformatics*. 2017;85(3):435-444. doi:10.1002/prot.25219

79. Desta IT, Porter KA, Xia B, Kozakov D, Vajda S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure*. 2020;28(9):1071-1081.e3. doi:10.1016/j.str.2020.06.006

80. Lek M, Karczewski KJ, Minikel E V., et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. Published online 2016. doi:10.1038/nature19057

81. Nolan LS, Chen J, Gonçalves AC, et al. Targeted deletion of the RNA-binding protein Caprin1 leads to progressive hearing loss and impairs recovery from noise exposure in mice. *Sci Rep*. 2022;12(1):2444. doi:10.1038/s41598-022-05657-2

82. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet*. Published online 2011. doi:10.1038/nrg3051

83. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

84. Chailangkarn T, Trujillo CA, Freitas BC, et al. A human neurodevelopmental model for Williams syndrome. *Nature*. 2016;536(7616):338-343. doi:10.1038/nature19067

85. Moujalled D, Strasser A, Liddell JR. Molecular mechanisms of cell death in neurological diseases. *Cell Death Differ*. 2021;28(7):2029-2044. doi:10.1038/s41418-021-00814-y

86. Calvo M, Villalobos C, Núñez L. Calcium Imaging in Neuron Cell Death. In: ; 2015:73-85. doi:10.1007/978-1-4939-2152-2_6

87. Demaurex N, Distelhorst C. Cell biology. Apoptosis--the calcium connection. *Science*. 2003;300(5616):65-67. doi:10.1126/science.1083628

88. Mincheva-Tasheva S, Obis E, Tamarit J, Ros J. Apoptotic cell death and altered calcium homeostasis caused by frataxin depletion in dorsal root ganglia neurons can be prevented by BH4 domain of Bcl-xL protein. *Hum Mol Genet*. 2014;23(7):1829-1841. doi:10.1093/hmg/ddt576

89. Sies H, Jones DP. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nat Rev Mol Cell Biol*. 2020;21(7):363-383. doi:10.1038/s41580-020-0230-3

90. Xu B, Chen S, Luo Y, et al. Calcium Signaling Is Involved in Cadmium-Induced Neuronal Apoptosis via Induction of Reactive Oxygen Species and Activation of MAPK/mTOR Network. *PLoS One*. 2011;6(4):e19052. doi:10.1371/journal.pone.0019052

91. Chen L, Xu B, Liu L, et al. Cadmium induction of reactive oxygen species activates the mTOR pathway, leading to neuronal cell death. *Free Radic Biol Med*. 2011;50(5):624-632. doi:10.1016/j.freeradbiomed.2010.12.032

92. Solomon S, Xu Y, Wang B, et al. Distinct Structural Features ofCaprin-1 Mediate Its Interaction with G3BP-1 and Its Induction of Phosphorylation of Eukaryotic Translation InitiationFactor 2α, Entry to Cytoplasmic Stress Granules, and Selective Interaction with a Subset of mRNAs. *Mol Cell Biol*. 2007;27(6):2324-2342. doi:10.1128/MCB.02300-06

93. De Rubeis S, Pasciuto E, Li KW, et al. CYFIP1 Coordinates mRNA Translation and Cytoskeleton Remodeling to Ensure Proper Dendritic Spine Formation. *Neuron*. 2013;79(6):1169-1182. doi:10.1016/j.neuron.2013.06.039

94. Darnell JC, Klann E. The translation of translational control by FMRP: therapeutic targets for FXS. *Nat Neurosci*. 2013;16(11):1530-1536. doi:10.1038/nn.3379

95. Mofatteh M. mRNA localization and local translation in neurons. *AIMS Neurosci*. 2020;7(3):299-310. doi:10.3934/Neuroscience.2020016

96. Velasco G, Francastel C. Genetics meets DNA methylation in rare diseases. *Clin Genet*. 2019;95(2):210-220. doi:10.1111/cge.13480

97. Kinoshita PF, Orellana AMM, Nakao VW, et al. The Janus face of ouabain in Na^+/K^+-ATPase and calcium signalling in neurons. *Br J Pharmacol*. 2022;179(8):1512-1524. doi:10.1111/bph.15419

98. Reich CG, Mason SE, Alger BE. Novel Form of LTD Induced by Transient, Partial Inhibition of the Na,K-Pump in Rat Hippocampal CA1 Cells. *J Neurophysiol*. 2004;91(1):239-247. doi:10.1152/jn.00722.2003

99. Pangrazzi L, Balasco L, Bozzi Y. Oxidative Stress and Immune System Dysfunction in Autism Spectrum Disorders. *Int J Mol Sci*. 2020;21(9):3293. doi:10.3390/ijms21093293

100. Curpan AS, Luca AC, Ciobica A. Potential Novel Therapies for Neurodevelopmental Diseases Targeting Oxidative Stress. *Oxid Med Cell Longev*. 2021;2021:1-13. doi:10.1155/2021/6640206

101. Avcil S, Uysal P, Avcil M, Alışık M, Biçer C. Dynamic thiol/disulfide homeostasis in children with attention deficit hyperactivity disorder and its relation with disease subtypes. *Compr Psychiatry*. 2017;73:53-60. doi:10.1016/j.comppsych.2016.11.003

102. Nakashima T, Sekiguchi T, Kuraoka A, et al. Molecular cloning of a human cDNA encoding a novel protein, DAD1, whose defect causes apoptotic cell death in hamster BHK21 cells. *Mol Cell Biol*. 1993;13(10):6367-6374. doi:10.1128/mcb.13.10.6367-6374.1993

103. Topf U, Suppanz I, Samluk L, et al. Quantitative proteomics identifies redox switches for global translation modulation by mitochondrially produced reactive oxygen species. *Nat Commun*. 2018;9(1):324. doi:10.1038/s41467-017-02694-8

104. Shenton D, Smirnova JB, Selley JN, et al. Global Translational Responses to Oxidative Stress Impact upon Multiple Levels of Protein Synthesis. *Journal of Biological Chemistry*. 2006;281(39):29011-29021. doi:10.1074/jbc.M601545200

105. Keithley EM. Pathology and mechanisms of cochlear aging. *J Neurosci Res*. 2020;98(9):1674-1684. doi:10.1002/jnr.24439

106. Hekselman I, Yeger-Lotem E. Mechanisms of tissue and cell-type specificity in heritable traits and diseases. *Nat Rev Genet*. 2020;21(3):137-150. doi:10.1038/s41576-019-0200-9

107. Shiina N, Yamaguchi K, Tokunaga M. RNG105 Deficiency Impairs the Dendritic Localization of mRNAs for Na+/K+ ATPase Subunit Isoforms and Leads to the Degeneration of Neuronal Networks. *Journal of Neuroscience*. 2010;30(38). doi:10.1523/JNEUROSCI.6386-09.2010

108. Lek M, Karczewski KJ, Minikel E v., et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291. doi:10.1038/nature19057

109. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat*. Published online May 22, 2019:humu.23798. doi:10.1002/humu.23798

110. Delle Vedove A, Natarajan J, Zanni G, et al. CAPRIN1P512L causes aberrant protein aggregation and associates with early-onset ataxia. *Cellular and Molecular Life Sciences*. 2022;79(10):526. doi:10.1007/s00018-022-04544-3

111. Gurovich Y, Hanani Y, Bar O, et al. Identifying facial phenotypes of genetic disorders using deep learning. *Nat Med*. Published online 2019. doi:10.1038/s41591-018-0279-0

112. Thelen MP, Wirth B, Kye MJ. Mitochondrial defects in the respiratory complex I contribute to impaired translational initiation via ROS and energy homeostasis in SMA motor neurons. *Acta Neuropathol Commun*. 2020;8(1):223. doi:10.1186/s40478-020-01101-6

113. Paoletti P, Bellone C, Zhou Q. NMDA receptor subunit diversity: impact on receptor properties, synaptic plasticity and disease. *Nat Rev Neurosci*. 2013;14(6):383-400. doi:10.1038/nrn3504

114. Bellone C, Nicoll RA. Rapid Bidirectional Switching of Synaptic NMDA Receptors. *Neuron*. 2007;55(5):779-785. doi:10.1016/j.neuron.2007.07.035

115. XiangWei W, Jiang Y, Yuan H. De novo mutations and rare variants occurring in NMDA receptors. *Curr Opin Physiol*. 2018;2:27-35. doi:10.1016/j.cophys.2017.12.013

116. Santos-Gómez A, Miguez-Cabello F, García-Recio A, et al. Disease-associated GRIN protein truncating variants trigger NMDA receptor loss-of-function. *Hum Mol Genet*. 2021;29(24):3859-3871. doi:10.1093/hmg/ddaa220

117. de Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 2014;515(7526):209-215. doi:10.1038/nature13772

118. Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216-221. doi:10.1038/nature13908

119. Endele S, Rosenberger G, Geider K, et al. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet*. 2010;42(11):1021-1026. doi:10.1038/ng.677

120. Carvill GL, Regan BM, Yendle SC, et al. GRIN2A mutations cause epilepsy-aphasia spectrum disorders. *Nat Genet*. 2013;45(9):1073-1076. doi:10.1038/ng.2727

121. Strehlow V, Heyne HO, Vlaskamp DRM, et al. *GRIN2A* -related disorders: genotype and functional consequence predict phenotype. *Brain*. 2019;142(1):80-92. doi:10.1093/brain/awy304

122. Liu XR, Xu XX, Lin SM, et al. GRIN2A Variants Associated With Idiopathic Generalized Epilepsies. *Front Mol Neurosci*. 2021;14. doi:10.3389/fnmol.2021.720984

123. Lemke JR, Lal D, Reinthaler EM, et al. Mutations in GRIN2A cause idiopathic focal epilepsy with rolandic spikes. *Nat Genet*. 2013;45(9):1067-1072. doi:10.1038/ng.2728

124. Lesca G, Rudolf G, Bruneau N, et al. GRIN2A mutations in acquired epileptic aphasia and related childhood focal epilepsies and encephalopathies with speech and language dysfunction. *Nat Genet*. 2013;45(9):1061-1066. doi:10.1038/ng.2726

125. Elmasri M, Hunter DW, Winchester G, et al. Common synaptic phenotypes arising from diverse mutations in the human NMDA receptor subunit GluN2A. *Commun Biol*. 2022;5(1):174. doi:10.1038/s42003-022-03115-3

126. Bading H. Nuclear calcium signalling in the regulation of brain function. *Nat Rev Neurosci*. 2013;14(9):593-608. doi:10.1038/nrn3531

127. Franchini L, Stanic J, Ponzoni L, et al. Linking NMDA Receptor Synaptic Retention to Synaptic Plasticity and Cognition. *iScience*. 2019;19:927-939. doi:10.1016/j.isci.2019.08.036

128. Strehlow V, Heyne HO, Vlaskamp DRM, et al. *GRIN2A* -related disorders: genotype and functional consequence predict phenotype. *Brain*. 2019;142(1):80-92. doi:10.1093/brain/awy304

129. Rodríguez-Palmero A, Boerrigter MM, Gómez-Andrés D, et al. DLG4-related synaptopathy: a new rare brain disorder. *Genetics in Medicine*. 2021;23(5):888-899. doi:10.1038/s41436-020-01075-9

130. Lelieveld SH, Reijnders MRF, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci*. 2016;19(9):1194-1196. doi:10.1038/nn.4352

131. Stanic J, Carta M, Eberini I, et al. Rabphilin 3A retains NMDA receptors at synaptic sites through interaction with GluN2A/PSD-95 complex. *Nat Commun*. 2015;6(1):10181. doi:10.1038/ncomms10181

132. Franchini L, Stanic J, Barzasi M, et al. Rabphilin-3A Drives Structural Modifications of Dendritic Spines Induced by Long-Term Potentiation. *Cells*. 2022;11(10):1616. doi:10.3390/cells11101616

133. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

134. Samocha KE, Robinson EB, Sanders SJ, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet*. 2014;46(9):944-950. doi:10.1038/ng.3050

135. Gaildrat P, Killian A, Martins A, Tournier I, Frébourg T, Tosi M. Use of Splicing Reporter Minigene Assay to Evaluate the Effect on Splicing of Unclassified Genetic Variants. In: ; 2010:249-257. doi:10.1007/978-1-60761-759-4_15

136. Montaville P, Schlicker C, Leonov A, Zweckstetter M, Sheldrick GM, Becker S. The C2A-C2B Linker Defines the High Affinity Ca2+ Binding Mode of Rabphilin-3A. *Journal of Biological Chemistry*. 2007;282(7):5015-5025. doi:10.1074/jbc.M606746200

137. Guillén J, Ferrer-Orta C, Buxaderas M, et al. Structural insights into the Ca^{2+} and $PI(4,5)P_2$ binding modes of the C2 domains of rabphilin 3A and synaptotagmin 1. *Proceedings of the National Academy of Sciences*. 2013;110(51):20503-20508. doi:10.1073/pnas.1316179110

138. Coudevylle N, Montaville P, Leonov A, Zweckstetter M, Becker S. Structural Determinants for Ca2+ and Phosphatidylinositol 4,5-Bisphosphate Binding by the C2A Domain of Rabphilin-3A. *Journal of Biological Chemistry*. 2008;283(51):35918-35928. doi:10.1074/jbc.M804094200

139. Gurgone A, Pizzo R, Raspanti A, et al. mGluR5 PAMs rescue cortical and behavioural defects in a mouse model of CDKL5 deficiency disorder. *Neuropsychopharmacology*. Published online August 9, 2022. doi:10.1038/s41386-022-01412-3

140. Marcantoni A, Cerullo MS, Buxeda P, et al. Amyloid Beta42 oligomers up-regulate the excitatory synapses by potentiating presynaptic release while impairing postsynaptic NMDA receptors. *J Physiol*. 2020;598(11):2183-2197. doi:10.1113/JP279345

141. Hardingham GE, Bading H. Synaptic versus extrasynaptic NMDA receptor signalling: implications for neurodegenerative disorders. *Nat Rev Neurosci*. 2010;11(10):682-696. doi:10.1038/nrn2911

142. Brozzi F, Diraison F, Lajus S, et al. Molecular Mechanism of Myosin Va Recruitment to Dense Core Secretory Granules. *Traffic*. 2012;13(1):54-69. doi:10.1111/j.1600-0854.2011.01301.x

143. Correia SS, Bassani S, Brown TC, et al. Motor protein–dependent transport of AMPA receptors into spines during long-term potentiation. *Nat Neurosci*. 2008;11(4):457-466. doi:10.1038/nn2063

144. Stanic J, Mellone M, Napolitano F, et al. Rabphilin 3A: A novel target for the treatment of levodopa-induced dyskinesias. *Neurobiol Dis*. 2017;108:54-64. doi:10.1016/j.nbd.2017.08.001

145. Maselli RA, Vázquez J, Schrumpf L, et al. Presynaptic congenital myasthenic syndrome with altered synaptic vesicle homeostasis linked to compound heterozygous sequence variants in *RPH3A*. *Mol Genet Genomic Med*. 2018;6(3):434-440. doi:10.1002/mgg3.370

146. Suzuki T, Koike Y, Ashikawa K, et al. Genome-wide association study of epilepsy in a Japanese population identified an associated region at chromosome 12q24. *Epilepsia*. 2021;62(6):1391-1400. doi:10.1111/epi.16911

147. Schlüter OM, Schnell E, Verhage M, et al. Rabphilin Knock-Out Mice Reveal That Rabphilin Is Not Required for Rab3 Function in Regulating Neurotransmitter Release. *The Journal of Neuroscience*. 1999;19(14):5834-5846. doi:10.1523/JNEUROSCI.19-14-05834.1999

148. Lee SB, Segura-Bayona S, Villamor-Payà M, et al. Tousled-like kinases stabilize replication forks and show synthetic lethality with checkpoint and PARP inhibitors. *Sci Adv*. Published online 2018. doi:10.1126/sciadv.aat4985

149. Segura-Bayona S, Stracker TH. The Tousled-like kinases regulate genome and epigenome stability: implications in development and disease. *Cellular and Molecular Life Sciences*. Published online 2019. doi:10.1007/s00018-019-03208-z

150. Bruinsma W, Berg J, Aprelia M, Medema RH. Tousled-like kinase 2 regulates recovery from a DNA damage-induced G2 arrest . *EMBO Rep*. Published online 2016. doi:10.15252/embr.201540767

151. Klimovskaia IM, Young C, Strømme CB, et al. Tousled-like kinases phosphorylate Asf1 to promote histone supply during DNA replication. *Nat Commun*. Published online 2014. doi:10.1038/ncomms4394

152. Segura-Bayona S, Knobel PA, Gonzalez-Buron H, et al. Differential requirements for Tousled-like kinases 1 and 2 in mammalian development. *Cell Death Differ*. Published online 2017. doi:10.1038/cdd.2017.108

153. Sandra Segura-Bayona, Marina Villamor-Payà, Camille Stephan-Otto Attolini, Travis H. Stracker. Tousled-like kinase activity is required for

transcriptional silencing and suppression of innate immune signaling. Published online 2019.

154. Töpf A, Oktay Y, Balaraju S, et al. Severe neurodevelopmental disease caused by a homozygous TLK2 variant. *European Journal of Human Genetics*. Published online 2019. doi:10.1038/s41431-019-0519-x

155. Mortuza GB, Hermida D, Pedersen AK, et al. Molecular basis of Tousled-Like Kinase 2 activation. *Nat Commun*. Published online 2018. doi:10.1038/s41467-018-04941-y

156. Töpf A, Oktay Y, Balaraju S, et al. Severe neurodevelopmental disease caused by a homozygous TLK2 variant. *European Journal of Human Genetics*. Published online 2019. doi:10.1038/s41431-019-0519-x

157. Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216-221. doi:10.1038/nature13908

158. O'Roak BJ, Deriziotis P, Lee C, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet*. Published online 2011. doi:10.1038/ng.835

159. Kopanos C, Tsiolkas V, Kouris A, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. Published online 2019. doi:10.1093/bioinformatics/bty897

160. Desmet FO, Hamroun D, Lalande M, Collod-Bëroud G, Claustres M, Béroud C. Human Splicing Finder: An online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. Published online 2009. doi:10.1093/nar/gkp215

161. Silva J, Aivio S, Knobel PA, et al. EXD2 governs germ stem cell homeostasis and lifespan by promoting mitoribosome integrity and translation. *Nat Cell Biol*. Published online 2018. doi:10.1038/s41556-017-0016-9

162. Reijnders MRF, Miller KA, Alvi M, et al. 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;102(6):1195-1203. doi:10.1016/j.ajhg.2018.04.014

163. Firth H V., Richards SM, Bevan AP, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet*. Published online 2009. doi:10.1016/j.ajhg.2009.03.010

164. Havrilla JM, Pedersen BS, Layer RM, Quinlan AR. A map of constrained coding regions in the human genome. *Nat Genet*. 2019;51(January). doi:10.1038/s41588-018-0294-6

165. Manning KS, Cooper TA. The roles of RNA processing in translating genotype to phenotype. *Nat Rev Mol Cell Biol*. Published online 2017. doi:10.1038/nrm.2016.139

166. Siwaszek A, Ukleja M, Dziembowski A. Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. *RNA Biol*. Published online 2014. doi:10.4161/rna.34406

167. Treiber DK, Shah NP. Ins and outs of kinase DFG motifs. *Chem Biol*. Published online 2013. doi:10.1016/j.chembiol.2013.06.001

168. Vijayan RSK, He P, Modi V, et al. Conformational analysis of the DFG-out kinase motif and biochemical profiling of structurally validated type II inhibitors. *J Med Chem*. Published online 2015. doi:10.1021/jm501603h

169. Roux KJ, Kim DI, Burke B, May DG. BioID: A Screen for Protein-Protein Interactions. *Curr Protoc Protein Sci*. 2018;91(1):19.23.1-19.23.15. doi:10.1002/cpps.51

170. Mortuza GB, Hermida D, Pedersen AK, et al. Molecular basis of Tousled-Like Kinase 2 activation. *Nat Commun*. Published online 2018. doi:10.1038/s41467-018-04941-y

171. Wessel SR, Mohni KN, Luzwick JW, Dungrawala H, Cortez D. Functional Analysis of the Replication Fork Proteome Identifies BET Proteins as PCNA Regulators. *Cell Rep*. Published online 2019. doi:10.1016/j.celrep.2019.08.051

172. Lalani SR, Safiullah AM, Fernbach SD, et al. Spectrum of CHD7 mutations in 110 individuals with CHARGE syndrome and genotype-phenotype correlation. *Am J Hum Genet*. Published online 2006. doi:10.1086/500273

173. Bernier R, Golzio C, Xiong B, et al. Disruptive CHD8 mutations define a subtype of autism early in development. *Cell*. Published online 2014. doi:10.1016/j.cell.2014.06.017

174. Manning BJ, Yusufzai T. The ATP-dependent chromatin remodeling enzymes CHD6, CHD7, and CHD8 exhibit distinct nucleosome binding and remodeling activities. *Journal of Biological Chemistry*. Published online 2017. doi:10.1074/jbc.M117.779470

175. Murawska M, Brehm A. CHD chromatin remodelers and the transcription cycle. *Transcription*. Published online 2011. doi:10.4161/trns.2.6.17840

176. Lelieveld SH, Reijnders MRF, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci*. 2016;19(9):1194-1196. doi:10.1038/nn.4352

177. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. Published online 2020. doi:10.1016/j.cell.2019.12.036

178. Sunavala-Dossabhoy G, Li Y, Williams B, De Benedetti A. A dominant negative mutant of TLK1 causes chromosome missegregation and aneuploidy in normal breast epithelial cells. *BMC Cell Biol*. Published online 2003. doi:10.1186/1471-2121-4-16

179. Liu H, Dowdle JA, Khurshid S, et al. Discovery of Stromal Regulatory Networks that Suppress Ras-Sensitized Epithelial Cell Proliferation. *Dev Cell*. Published online 2017. doi:10.1016/j.devcel.2017.04.024

180. Lombardo M V., Moon HM, Su J, Palmer TD, Courchesne E, Pramparo T. Maternal immune activation dysregulation of the fetal brain transcriptome and relevance to the pathophysiology of autism spectrum disorder. *Mol Psychiatry*. Published online 2018. doi:10.1038/mp.2017.15

181. Carpentier PA, Dingman AL, Palmer TD. Placental TNF-α signaling in illness-induced complications of pregnancy. *American Journal of Pathology*. Published online 2011. doi:10.1016/j.ajpath.2011.02.042

182. Taylor AMR. Chromosome instability syndromes. *Best Pract Res Clin Haematol*. Published online 2001. doi:10.1053/beha.2001.0158

183. Waltes R, Kalb R, Gatei M, et al. Human RAD50 Deficiency in a Nijmegen Breakage Syndrome-like Disorder. *Am J Hum Genet*. Published online 2009. doi:10.1016/j.ajhg.2009.04.010

184. Mi W, Guan H, Lyu J, et al. YEATS2 links histone acetylation to tumorigenesis of non-small cell lung cancer. *Nat Commun*. Published online 2017. doi:10.1038/s41467-017-01173-4

185. Zhao D, Guan H, Zhao S, et al. YEATS2 is a selective histone crotonylation reader. *Cell Res*. Published online 2016. doi:10.1038/cr.2016.49

186. Yeetong P, Pongpanich M, Srichomthong C, et al. TTTCA repeat insertions in an intron of YEATS2 in benign adult familial myoclonic epilepsy type 4. *Brain*. Published online 2019. doi:10.1093/brain/awz267

187. Tranfaglia MR, Thibodeaux C, Mason DJ, et al. Repurposing available drugs for neurodevelopmental disorders: The fragile X experience. *Neuropharmacology*. Published online 2019. doi:10.1016/j.neuropharm.2018.05.004

188. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. Published online 2019. doi:10.1093/nar/gky1131

189. Oughtred R, Stark C, Breitkreutz BJ, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res*. Published online 2019. doi:10.1093/nar/gky1079

190. Knight JDR, Choi H, Gupta GD, et al. ProHits-viz: A suite of web tools for visualizing interaction proteomics data. *Nat Methods*. Published online 2017. doi:10.1038/nmeth.4330

191. Pavinato L, Nematian-Ardestani E, Zonta A, et al. KCNK18 Biallelic Variants Associated with Intellectual Disability and Neurodevelopmental Disorders Alter TRESK Channel Activity. *Int J Mol Sci*. 2021;22(11):6064. doi:10.3390/ijms22116064

192. Achour A, Koopmann T, Castel R, et al. A new gene associated with a β-thalassemia phenotype: the observation of variants in SUPT5H. *Blood*. 2020;136(15):1789-1793. doi:10.1182/blood.2020005934

193. Nguengang Wakap S, Lambert DM, Olry A, et al. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *European Journal of Human Genetics*. Published online 2020. doi:10.1038/s41431-019-0508-0

194. Hennekam RCM. Care for patients with ultra-rare disorders. *Eur J Med Genet*. Published online 2011. doi:10.1016/j.ejmg.2010.12.001

195. Skraban CM, Wells CF, Markose P, et al. WDR26 Haploinsufficiency Causes a Recognizable Syndrome of Intellectual Disability, Seizures, Abnormal Gait, and Distinctive Facial Features. *Am J Hum Genet*. Published online 2017. doi:10.1016/j.ajhg.2017.06.002

196. Clevers H, Nusse R. Wnt/β-catenin signaling and disease. *Cell*. Published online 2012. doi:10.1016/j.cell.2012.05.012

197. Ye Y, Tang X, Sun Z, Chen S. Upregulated WDR26 serves as a scaffold to coordinate PI3K/AKT pathway-driven breast cancer cell growth, migration, and invasion. *Oncotarget*. 2016;7(14):17854-17869. doi:10.18632/oncotarget.7439

198. Jiao Q, Sun H, Zhang H, et al. The combination of whole-exome sequencing and copy number variation sequencing enables the diagnosis of rare neurological disorders. *Clin Genet*. 2019;96(2):140-150. doi:10.1111/cge.13548

199. Takata A, Miyake N, Tsurusaki Y, et al. Integrative Analyses of De Novo Mutations Provide Deeper Biological Insights into Autism Spectrum Disorder. *Cell Rep*. 2018;22(3):734-747. doi:10.1016/j.celrep.2017.12.074

200. Jagadeesh KA, Wenger AM, Berger MJ, et al. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet*. Published online 2016. doi:10.1038/ng.3703

201. Sadleir LG, Scheffer IE, Smith S, Carstensen B, Farrell K, Connolly MB. EEG features of absence seizures in idiopathic generalized epilepsy: Impact of syndrome, age, and state. *Epilepsia*. Published online 2009. doi:10.1111/j.1528-1167.2008.02001.x

202. Griffith EM, Pennington BF, Wehner EA, Rogers SJ. Executive functions in young children with autism. *Child Dev*. Published online 1999. doi:10.1111/1467-8624.00059

203. Keller R, Chieregato S, Bari S, et al. Autism in adulthood: Clinical and demographic characteristics of a cohort of five hundred persons with autism analyzed by a novel multistep network model. *Brain Sci*. Published online 2020. doi:10.3390/brainsci10070416

204. Crunelli V, Leresche N. Childhood absence epilepsy: Genes, channels, neurons and networks. *Nat Rev Neurosci*. Published online 2002. doi:10.1038/nrn811

205. Cope DW, Di Giovanni G, Fyson SJ, et al. Enhanced tonic GABA A inhibition in typical absence epilepsy. *Nat Med*. Published online 2009. doi:10.1038/nm.2058

206. Courtiol E, Wilson DA. The olfactory thalamus: Unanswered questions about the role of the mediodorsal thalamic nucleus in olfaction. *Front Neural Circuits*. Published online 2015. doi:10.3389/fncir.2015.00049

207. Sorokin JM, Davidson TJ, Frechette E, et al. Bidirectional Control of Generalized Epilepsy Networks via Rapid Real-Time Switching of Firing Mode. *Neuron*. Published online 2017. doi:10.1016/j.neuron.2016.11.026

208. Fogerson PM, Huguenard JR. Tapping the Brakes: Cellular and Synaptic Mechanisms that Regulate Thalamic Oscillations. *Neuron*. Published online 2016. doi:10.1016/j.neuron.2016.10.024

209. Paz JT, Davidson TJ, Frechette ES, et al. Closed-loop optogenetic control of thalamus as a tool for interrupting seizures after cortical injury. *Nat Neurosci*. Published online 2013. doi:10.1038/nn.3269

210. Hofman J, Hutny M, Sztuba K, Paprocka J. Corpus callosum agenesis: An insight into the etiology and spectrum of symptoms. *Brain Sci*. Published online 2020. doi:10.3390/brainsci10090625

211. Edwards TJ, Sherr EH, Barkovich AJ, Richards LJ. Clinical, genetic and imaging findings identify new causes for corpus callosum development syndromes. *Brain*. Published online 2014. doi:10.1093/brain/awt358

212. Rosenfeld JA, Lacassie Y, El-Khechen D, et al. New cases and refinement of the critical region in the 1q41q42 microdeletion syndrome. *Eur J Med Genet*. Published online 2011. doi:10.1016/j.ejmg.2010.10.002

213. Stirnimann CU, Petsalaki E, Russell RB, Müller CW. WD40 proteins propel cellular networks. *Trends Biochem Sci*. Published online 2010. doi:10.1016/j.tibs.2010.04.003

214. Goto T, Matsuzawa J, Iemura SI, Natsume T, Shibuya H. WDR26 is a new partner of Axin1 in the canonical Wnt signaling pathway. *FEBS Lett*. Published online 2016. doi:10.1002/1873-3468.12180

215. Wang L, Zhou K, Fu Z, et al. Brain Development and Akt Signaling: the Crossroads of Signaling Pathway and Neurodevelopmental Diseases. *Journal of Molecular Neuroscience*. Published online 2017. doi:10.1007/s12031-016-0872-y

216. Gada K, Plant LD. Two-pore domain potassium channels: emerging targets for novel analgesic drugs: IUPHAR Review 26. *Br J Pharmacol*. Published online 2019. doi:10.1111/bph.14518

217. Andres-Enguix I, Shang L, Stansfeld PJ, et al. Functional analysis of missense variants in the TRESK (KCNK18) K + channel. *Sci Rep*. Published online 2012. doi:10.1038/srep00237

218. Enyedi P, Czirják G. Properties, regulation, pharmacology, and functions of the K2P channel, TRESK. *Pflugers Arch*. Published online 2015. doi:10.1007/s00424-014-1634-8

219. Czirják G, Enyedi P. TRESK background K+ channel is inhibited by phosphorylation via two distinct pathways. *Journal of Biological Chemistry*. Published online 2010. doi:10.1074/jbc.M110.102020

220. Czirják G, Tóth ZE, Enyedi P. The Two-pore Domain K+ Channel, TRESK, Is Activated by the Cytoplasmic Calcium Signal through Calcineurin. *Journal of Biological Chemistry*. Published online 2004. doi:10.1074/jbc.M312229200

221. Bautista DM, Sigal YM, Milstein AD, et al. Pungent agents from Szechuan peppers excite sensory neurons by inhibiting two-pore potassium channels. *Nat Neurosci*. Published online 2008. doi:10.1038/nn.2143

222. Yoo SH, Liu J, Sabbadini M, Au P, Xie G xi, Yost CS. Regional expression of the anesthetic-activated potassium channel TRESK in the rat nervous system. *Neurosci Lett*. Published online 2009. doi:10.1016/j.neulet.2009.08.062

223. Pettingill P, Weir GA, Wei T, et al. A causal role for TRESK loss of function in migraine mechanisms. *Brain*. Published online 2019. doi:10.1093/brain/awz342

224. Royal P, Andres-Bilbe A, Ávalos Prado P, et al. Migraine-Associated TRESK Mutations Increase Neuronal Excitability through Alternative Translation Initiation and Inhibition of TREK. *Neuron*. Published online 2019. doi:10.1016/j.neuron.2018.11.039

225. Lafrenière RG, Cader MZ, Poulin JF, et al. A dominant-negative mutation in the TRESK potassium channel is linked to familial migraine with aura. *Nat Med*. Published online 2010. doi:10.1038/nm.2216

226. Rainero I, Rubino E, Gallone S, et al. KCNK18 (TRESK) genetic variants in Italian patients with migraine. *Headache*. Published online 2014. doi:10.1111/head.12439

227. Han JY, Jang JH, Park J, Lee IG. Targeted next-generation sequencing of Korean patients with developmental delay and/or intellectual disability. *Front Pediatr*. 2018;6(December):1-9. doi:10.3389/fped.2018.00391

228. Imbrici P, Nematian-Ardestani E, Hasan S, Pessia M, Tucker SJ, D'Adamo MC. Altered functional properties of a missense variant in the TRESK K+ channel

(KCNK18) associated with migraine and intellectual disability. *Pflugers Arch*. Published online 2020. doi:10.1007/s00424-020-02382-5

229. Schalock RL, Luckasson R, Tassé MJ. The contemporary view of intellectual and developmental disabilities: Implications for psychologists. *Psicothema*. Published online 2019. doi:10.7334/psicothema2019.119

230. De Giorgio A. The roles of motor activity and environmental enrichment in intellectual disability. *Somatosens Mot Res*. Published online 2017. doi:10.1080/08990220.2016.1278204

231. Czirják G, Enyedi P. Targeting of calcineurin to an NFAT-like docking site is required for the calcium-dependent activation of the background K+ channel, TRESK. *Journal of Biological Chemistry*. Published online 2006. doi:10.1074/jbc.M602495200

232. Yoshida S, Plant S. Mechanism of release of Ca2+ from intracellular stores in response to ionomycin in oocytes of the frog Xenopus laevis. *J Physiol*. Published online 1992. doi:10.1113/jphysiol.1992.sp019419

233. Cadaveira-Mosquera A, Pérez M, Reboreda A, Rivas-Ramírez P, Fernández-Fernández D, Lamas JA. Expression of K2P channels in sensory and motor neurons of the autonomic nervous system. *Journal of Molecular Neuroscience*. Published online 2012. doi:10.1007/s12031-012-9780-y

234. Kang D, Kim D. TREK-2 (K2P10.1) and TRESK (K2P18.1) are major background K+ channels in dorsal root ganglion neurons. *Am J Physiol Cell Physiol*. Published online 2006. doi:10.1152/ajpcell.00629.2005

235. Maher BH, Taylor M, Stuart S, et al. Analysis of 3 common polymorphisms in the KCNK18 gene in an Australian Migraine Case-control cohort. *Gene*. Published online 2013. doi:10.1016/j.gene.2013.07.030

236. Czirják G, Vuity D, Enyedi P. Phosphorylation-dependent binding of 14-3-3 proteins controls TRESK regulation. *Journal of Biological Chemistry*. Published online 2008. doi:10.1074/jbc.M800712200

237. Rahm AK, Gierten J, Kisselbach J, et al. PKC-dependent activation of human K 2P18.1 K + channels. *Br J Pharmacol*. 2012;166(2):764-773. doi:10.1111/j.1476-5381.2011.01813.x

238. Sehgal SA, Hassan M, Rashid S. Pharmacoinformatics elucidation of potential drug targets against migraine to target ion channel protein KCNK18. *Drug Des Devel Ther*. Published online 2014. doi:10.2147/DDDT.S63096

239. Lengyel M, Dobolyi A, Czirják G, Enyedi P. Selective and state-dependent activation of TRESK (K2P18.1) background potassium channel by cloxyquin. *Br J Pharmacol*. Published online 2017. doi:10.1111/bph.13821

240. Czirják G, Vuity D, Enyedi P. Phosphorylation-dependent binding of 14-3-3 proteins controls TRESK regulation. *Journal of Biological Chemistry*. Published online 2008. doi:10.1074/jbc.M800712200

241. Cornell B, Toyo-oka K. 14-3-3 proteins in brain development: Neurogenesis, neuronal migration and neuromorphogenesis. *Front Mol Neurosci*. Published online 2017. doi:10.3389/fnmol.2017.00318

242. Guella I, McKenzie MB, Evans DM, et al. De Novo Mutations in YWHAG Cause Early-Onset Epilepsy. *Am J Hum Genet*. Published online 2017. doi:10.1016/j.ajhg.2017.07.004

243. Guella I, McKenzie MB, Evans DM, et al. De Novo Mutations in YWHAG Cause Early-Onset Epilepsy. *Am J Hum Genet*. Published online 2017. doi:10.1016/j.ajhg.2017.07.004

244. Sano Y, Inamura K, Miyake A, et al. A novel two-pore domain K+ channel, TRESK, is localized in the spinal cord. *Journal of Biological Chemistry*. Published online 2003. doi:10.1074/jbc.M206810200

245. Liu C, Au JD, Zou HL, Cotten JF, Yost CS. Potent activation of the human tandem pore domain K channel TRESK with clinical concentrations of volatile anesthetics. *Anesth Analg*. Published online 2004. doi:10.1213/01.ANE.0000136849.07384.44

246. Yi C, Spitters TW, Al-Far EADA, et al. A calcineurin-mediated scaling mechanism that controls a K+-leak channel to regulate morphogen and growth factor transcription. *bioRxiv*. Published online 2021. doi:https://doi.org/10.1101/2020.07.03.186403

247. Sun C, Zou M, Li L, et al. Association study between inwardly rectifying potassium channels 2.1 and 4.1 and autism spectrum disorders. *Life Sci*. Published online 2018. doi:10.1016/j.lfs.2018.10.012

248. Sicca F, Ambrosini E, Marchese M, et al. Gain-of-function defects of astrocytic Kir4.1 channels in children with autism spectrum disorders and epilepsy. *Sci Rep*. Published online 2016. doi:10.1038/srep34325

249. Sicca F, Imbrici P, D'Adamo MC, et al. Autism with Seizures and Intellectual Disability: Possible Causative Role of Gain-of-function of the Inwardly-

Rectifying K + Channel Kir4.1. *Neurobiol Dis*. Published online 2011. doi:10.1016/j.nbd.2011.03.016

250. D'Adamo MC, Catacuzzeno L, di Giovanni G, Franciolini F, Pessia M. K+ channelepsy: Progress in the neurobiology of potassium channels and epilepsy. *Front Cell Neurosci*. Published online 2013. doi:10.3389/fncel.2013.00134

251. Hasan S, Balobaid A, Grottesi A, et al. Lethal digenic mutations in the K+ channels kir4.1 (KCNJ10) and SLACK (KCNT1) associated with severe-disabling seizures and neurodevelopmental delay. *J Neurophysiol*. Published online 2017. doi:10.1152/jn.00284.2017

252. Luca G, Ilenio S, Martino C, et al. Update on the implication of potassium channels in autism: K+ channelautism spectrum disorder. *Front Cell Neurosci*. Published online 2015. doi:10.3389/fncel.2015.00034

253. D'Adamo MC, Liantonio A, Conte E, Pessia M, Imbrici P. Ion Channels Involvement in Neurodevelopmental Disorders. *Neuroscience*. Published online 2020. doi:10.1016/j.neuroscience.2020.05.032

254. Kang D, Kim GT, Kim EJ, et al. Lamotrigine inhibits TRESK regulated by G-protein coupled receptor agonists. *Biochem Biophys Res Commun*. Published online 2008. doi:10.1016/j.bbrc.2008.01.008

255. Czirják G, Enyedi P. The LQLP calcineurin docking site is a major determinant of the calcium-dependent activation of human TRESK background K+ channel. *Journal of Biological Chemistry*. Published online 2014. doi:10.1074/jbc.M114.577684

256. Chen X, Qi Y, Wu Z, et al. Structural insights into preinitiation complex assembly on core promoters. *Science (1979)*. 2021;372(6541). doi:10.1126/science.aba8490

257. Core L, Adelman K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev*. 2019;33(15-16):960-982. doi:10.1101/gad.325142.119

258. Cramer P. Organization and regulation of gene transcription. *Nature*. 2019;573(7772):45-54. doi:10.1038/s41586-019-1517-4

259. Jonkers I, Kwak H, Lis JT. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife*. 2014;3. doi:10.7554/eLife.02407

260. Hu S, Peng L, Xu C, Wang Z, Song A, Chen FX. SPT5 stabilizes RNA polymerase II, orchestrates transcription cycles, and maintains the enhancer landscape. *Mol Cell*. 2021;81(21):4425-4439.e6. doi:10.1016/j.molcel.2021.08.029

261. Lou J, Ye Y, Sun M, Zhao Y, Fu Y, Liu Y. A stepwise haematological screening and whole-exome sequencing reveal multiple mutations from SUPT5H causing an elevation of Hb A2 from a cohort of 47336 individuals. *Int J Lab Hematol*. Published online 2022. doi:10.1111/ijlh.13959

262. Stachora AA, Enríquez Schäfer R, Pohlmeier M, Maier G, Ponstingl H. Human Supt5h protein, a putative modulator of chromatin structure, is reversibly phosphorylated in mitosis. *FEBS Lett*. 1997;409(1):74-78. doi:10.1016/S0014-5793(97)00486-9

263. Rapaport F, Boisson B, Gregor A, et al. Negative selection on human genes underlying inborn errors depends on disease outcome and both the mode and mechanism of inheritance. *Proceedings of the National Academy of Sciences*. 2021;118(3). doi:10.1073/pnas.2001248118

264. Garel C, Cont I, Alberti C, Josserand E, Moutard ML, Ducou le Pointe H. Biometry of the Corpus Callosum in Children: MR Imaging Reference Data. *American Journal of Neuroradiology*. 2011;32(8):1436-1443. doi:10.3174/ajnr.A2542

265. Nguyen S, Kannan M, Gaborit M, Collins SC, Yalcin B. Quantitative Neuroanatomical Phenotyping of the Embryonic Mouse Brain. *Curr Protoc*. 2022;2(7). doi:10.1002/cpz1.509

266. Warren Z, McPheeters ML, Sathe N, Foss-Feig JH, Glasser A, Veenstra-VanderWeele J. A Systematic Review of Early Intensive Intervention for Autism Spectrum Disorders. *Pediatrics*. 2011;127(5):e1303-e1311. doi:10.1542/peds.2011-0426

267. Lappé M, Lau L, Dudovitz RN, Nelson BB, Karp EA, Kuo AA. The Diagnostic Odyssey of Autism Spectrum Disorder. *Pediatrics*. 2018;141(Supplement 4):S272-S279. doi:10.1542/peds.2016-4300C

268. Henderson LB, Applegate CD, Wohler E, Sheridan MB, Hoover-Fong J, Batista DAS. The impact of chromosomal microarray on clinical management: a retrospective analysis. *Genet Med*. 2014;16(9):657-664. doi:10.1038/gim.2014.18

269. Dixon-Salazar TJ, Silhavy JL, Udpa N, et al. Exome sequencing can improve diagnosis and alter patient management. *Sci Transl Med*. 2012;4(138):138ra78. doi:10.1126/scitranslmed.3003544

270. Wolff M, Johannesen KM, Hedrich UBS, et al. Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. *Brain*. 2017;140(5):1316-1336. doi:10.1093/brain/awx054

271. Sahin M, Sur M. Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. *Science (1979)*. 2015;350(6263). doi:10.1126/science.aab3897

272. Hiatt SM, Lawlor JMJ, Handley LH, et al. Long-read genome sequencing for the molecular diagnosis of neurodevelopmental disorders. *Human Genetics and Genomics Advances*. 2021;2(2):100023. doi:10.1016/j.xhgg.2021.100023

273. Conlin LK, Aref-Eshghi E, McEldrew DA, Luo M, Rajagopalan R. Long-read sequencing for molecular diagnostics in constitutional genetic disorders. *Hum Mutat*. 2022;43(11):1531-1544. doi:10.1002/humu.24465

274. Noyes MD, Harvey WT, Porubsky D, et al. Familial long-read sequencing increases yield of de novo mutations. *The American Journal of Human Genetics*. 2022;109(4):631-646. doi:10.1016/j.ajhg.2022.02.014

275. Bertoli-Avella AM, Kandaswamy KK, Khan S, et al. Combining exome/genome sequencing with data repository analysis reveals novel gene-disease associations for a wide range of genetic disorders. *Genet Med*. 2021;23(8):1551-1568. doi:10.1038/s41436-021-01159-0

276. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26. doi:10.1038/nbt.1754

277. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22(22):4673-4680. doi:10.1093/nar/22.22.4673

278. Piccoli G, Verpelli C, Tonna N, et al. Proteomic Analysis of Activity-Dependent Synaptic Plasticity in Hippocampal Neurons. *J Proteome Res*. 2007;6(8):3203-3215. doi:10.1021/pr0701308

279. Guerois R, Nielsen JE, Serrano L. Predicting Changes in the Stability of Proteins and Protein Complexes: A Study of More Than 1000 Mutations. *J Mol Biol*. 2002;320(2):369-387. doi:10.1016/S0022-2836(02)00442-4

Personal scientific production

https://orcid.org/0000-0002-7630-8365

1. *Rare coding variation illuminates the allelic architecture, risk genes, cellular expression patterns, and phenotypic context of autism* | J. M. Fu, F. K. Satterstrom, M. Peng, H. Brand, R. L. Collins, S. Dong, B. Wamsley, L. Klei, L. Wang, S. P. Hao, C. R. Stevens, C. Cusick, M. Babadi, E. Banks, B. Collins, S. Dodge, S. B. Gabriel, L. Gauthier, S. K. Lee12 , L. L. A. Ljungdahl, B. Mahjani, L. Sloofman, A. Smirnov, M. Barbosa, C. Betancur, A. Brusco, B. H. Y. Chung, E. H. Cook, M. L. Cuccaro, E. Domenici, G. B. Ferrero, J. J. Gargus, G. E. Herman, I. Hertz-Picciotto, P. Maciel, D. S. Manoach, M. R. Passos Bueno, A. M. Persico, A. Renieri, J. S. Sutcliffe, F. Tassone, E. Trabetti, G. Campos, S. Cardaropoli, D. Carli, M. C.Y. Chan, C. Fallerini, E. Giorgio, A. C. Girard, E. Hansen Kiss, S. Lun Lee, C. Lintas, Y. Ludena, R. Nguyen, **L. Pavinato**, M. Pericak-Vance, I. Pessah, R. Schmidt, M. Smith, C. I. C. Souza, S. Trajkova, J. Y. T. Wang, M. H. C. Yu, The Autism Sequencing Consortium (ASC), Broad Institute Center for Common Disease Genomics (Broad-CCDG) , iPSYCH-BROAD Consortium, D. J. Cutler, S. De Rubeis, J. D. Buxbaum, M. J. Daly, B. Devlin, K. Roeder, S. J. Sanders, M. E. Talkowski
Nature Genetics | 2022-08-18| journal-article. https://doi.org/10.1038/s41588-022-01104-0

2. *CAPRIN1* haploinsufficiency causes a neurodevelopmental disorder with language impairment, ADHD and ASD | **L. Pavinato**, A. Delle Vedove, D. Carli, M. Ferrero, S. Carestiato, J. L. Howe, E. Agolini, D. A. Coviello, I. van de Laar, P. Y. Billie Au, E. Di Gregorio, A. Fabbiani, S. Croci, M. A. Mencarelli, L. P. Bruno, A. Renieri, D. Veltra, C. Sofocleous, L. Faivre, B. Mazel, H. Safraou, A-S. Denommé-Pichon, M. A. van Slegtenhorst, N. Giesbertz, R. H. van Jaarsveld, A. Childers, R. C. Rogers, A. Novelli, S. De Rubeis, J. D. Buxbaum, S. W. Scherer, G. B. Ferrero, B. Wirth, A. Brusco
Brain | 2022-07-27 | journal-article. https://doi.org/10.1093/brain/awac278

3. p140Cap regulates the composition and localization of the NMDAR complex in synaptic lipid rafts | C. Angelini, A. Morellato, A. Alfieri, **L. Pavinato**, T. Cravero, O. T. Bianciotto, V. Salemme, D. Natalini, G., A. Raspanti, T. Garofalo, D. Valdembri, G. Serini, A. Marcantoni, A. Becchetti, M. Giustetto, E. Turco, P. Defilippi.
The Journal of Neuroscience |2022-08-11 | journal-article
https://doi.org/10.1523/JNEUROSCI.1775-21.2022

4. *Truncating variants in NOVA2 disrupt alternative splicing events and lead to a heterogeneous neurodevelopmental disorder with variable clinical course* | M. Scala, N. Drouot, S. C. MacLennan, M. W. Wessels, M. Krygier, **L. Pavinato**, A. Telegrafi, S. A. de Man, M. van Slegtenhorst, M. Iacomino, F. Madia, P. Scudieri, T. Giacomini, G. Nobile, M. M. Mancardi, G. Balagura, G. B. Ferrero, G. Galloni, A. Verrotti, M. Umair, A. Khan, J. Liebelt, A. Brusco, B. S. Lipska Ziętkiewicz, J. J. Saris, N. Charlet-Berguerand, F. Zara, P. Striano, A. Piton

Human Mutation. 2022-09 |journal-article. https://doi.org/10.1002/humu.24414

5. The Emerging Roles of Long Non-Coding RNAs in Intellectual Disability and Related Neurodevelopmental Disorders | C. Liaci, L. Prandi, **L. Pavinato**, A. Brusco, M. Maldotti, I. Molineris, S. Oliviero, G. R. Merlo

International Journal of Molecular Sciences | review. doi.org/10.3390/ijms23116118

6. *Drosophila functional screening of de novo variants in autism uncovers deleterious variants and facilitates discovery of rare neurodevelopmental diseases.* | P. C. Marcogliese, S.L. Deal, J. Andrews, J. M. Harnish, V. H. Bhavana, H. K. Graves, S. Jangam, X. Luo, N. Liu, D. Bei, Y-H. Chao, B. Hull, P-T Lee, H. Pan, P. Bhadane, M-C Huang, C. M. Longley, H-T Chao, H-l Chung, N. A. Haelterman, O. Kanca, S. N. Manivannan, L. Z. Rossetti, R. J. German, A. Gerard, E. M. C. Schwaibold, S. Fehr, R. Guerrini, A. Vetro, E. England, C. N. Murali, T. S. Barakat, M. F. van Dooren, M. Wilke, M. van Slegtenhorst, G. Lesca, I. Sabatier, N Chatron, C. A. Brownstein, J. A. Madden, P. B. Agrawal, B. Keren, T. Courtin, L. Perrin, M. Brugger, T. Roser, S. Leiz, F. T. Mau-Them, J. Delanne, E. Sukarova-Angelovska, S. Trajkova, E. Rosenhahn, V. Strehlow, K. Platzer,R. Keller, **L. Pavinato**, A. Brusco, J. A. Rosenfeld, R. Marom, M. F. Wangler, S. Yamamoto

Cell Reports. 2022-03-15 | journal-article. DOI: 10.1016/j.celrep.2022.110517

7. *Functional analysis of TLK2 variants and their proximal interactomes implicates impaired kinase activity and chromatin maintenance defects in their pathogenesis.* | **L. Pavinato**, M. Villamor-Payà, M. Sanchiz-Calvo, C. Andreoli, M. Gay, M. Vilaseca, G. Arauz-Garofalo, T. Pippucci, V. Prota, D. Carli, E. Giogio, F. C. Radio, V. Antona, M. Giuffrè, K. Ranguin, C. Colson, S. De Rubeis, P. Dimartino, J. D. Buxbaum, G. B. Ferrero, M. Tartaglia, S. Martinelli, T. H. Stracker, A. Brusco

Journal of Medical Genetics. 2022-02 | journal-article. DOI: 10.1136/jmedgenet-2020-107281

8. *Biallelic PI4KA variants cause a novel neurodevelopmental syndrome with hypomyelinating leukodystrophy.* |E. Verdura, A. Rodríguez-Palmero, V. Vélez-

Santamaria, L. Planas-Serra, I. de la Calle, M. Raspall-Chaure, A. Roubertie, M. Benkirane, F. Saettini, **L. Pavinato**, G. Mandrile, M. O'Leary, E. O'Heir, E. Barredo, A. Chacón, V. Michaud, C. Goizet, M. Ruiz, A. Schlüter, I. Rouvet, J. Sala-Coromina, C. Fossati, M. Iascone, F. Canonico, A. Marcé-Grau, P. de Souza, D. R. Adams, C. Casasnovas, H. L. Rehm, H. C Mefford, L. González Gutierrez-Solana, A. Brusco, M. Koenig, A. Macaya, A. Pujol

Brain | 2021-08-20 | journal-article. DOI: 10.1093/brain/awab124

9. *KCNK18 biallelic variants associated with intellectual disability and autism spectrum disorder alter TRESK channel activity.* | **L. Pavinato**, E. Nematian-Ardestani, A. Zonta, S. De Rubeis, J. D. Buxbaum, C. Mancini, A. Bruselles, M. Tartaglia, M. Pessia, S. J. Tucker, M. C. D'Adamo, A. Brusco

International Journal of Molecular Sciences. 2021-06-04 | journal-article. DOI: 10.3390/ijms22116064

10. Expanding the clinical phenotype of the ultra-rare Skraban-Deardorff syndrome: Two novel individuals with *WDR26* loss-of-function variants and a literature review. | **L. Pavinato**, S. Trajkova, E. Grosso, E. Giorgio, A. Bruselles, F. C. Radio, T. Pippucci, P. Dimartino, M. Tartaglia, A. Petlichkovski, S. De Rubeis, J. D. Buxbaum, G. B. Ferrero, R. Keller, A. Brusco

American Journal of Medical Genetics Part A. 2021-03-06 | journal-article. DOI: 10.1002/ajmg.a.62157

11. *New Insights into Potocki-Shaffer Syndrome: Report of Two Novel Cases and Literature Review.* | S. Trajkova, E. Di Gregorio, G. B. Ferrero, D. Carli, **L. Pavinato**, G. Delplancq, P. Kuentz, A. Brusco

Brain Sciences | 2020-10-28 | journal-article. DOI: 10.3390/brainsci10110788

12. *Quando il disturbo del neurosviluppo ha un substrato genetico: un caso di sindrome di Kleefstra.* | L. Sorasio, L. Franceschi, **L. Pavinato**, A. Peduto.

Medico e Bambino, Volume XXIV, Aprile 2021, Numero 4. | caso contributivo. DOI: https://doi.org/10.53126/MEBXXIV114

Acknowledgements

Realizing you are ending your PhD is satisfactory and confusing at the same time, it is the conclusion of your career as student and the starting of the one as scientist, giving raise to new doubts and fears.

Life in the laboratory can be very hard sometimes, with too many deadlines, failed experiments and unexplainable scientific phenomena. But the feeling when an experiment works, when a diagnosis is identified thanks to your efforts or when you discover something new is priceless, and it is the fuel makes you hold on even when you want to quit, and that feeds your curiosity.

Thank all the people that have contributed to my PhD journey would be impossible, but I want to recognize at least some of them.

The first thank is for my mentor and supervisor, Professor Alfredo Brusco, who has always supported my work and patiently listened to me whenever I needed his help. Once I heard that it is important to choose a mentor that supports and guides you, giving you at the same time the possibility to make mistakes and understand: at the end of my PhD I can say that I agree with that suggestion.

Other people that have been very important are my colleagues: Marta Ferrero and Elisa Pozzi, -who are the ones that welcomed me in the lab-, Simona Cardaropoli and Slavica Trajkova, who are the remaining part of the “old group” with me, and the people with whom I shared the most moments and experiences. And then the new ones, Verdiana Pullano, Chiara Giovenino, Silvia Carestiato and Antonina Rinninella. Each one of them has been important for me, and I will always thank them for being able to being able to stand by me gently during some difficult times, and for bringing smiles and good vibes every day at work.

Another important piece of my PhD journey has been my stay in Germany, and I will be always grateful to Professor Brunhilde Wirth for this opportunity. My research period in Germany has been such an incredible time of growth for my career, and the tremendous support and admiration Brunhilde has always shown towards me have been very important for me.

All the members of Brunhilde’s lab will always have a place in my heart, for being such supportive and kind with me from the very first day. I want to particularly thank my favourite lab mates, Eleonora Zilio, Andrea Delle Vedove and Tamas Schmidt: thanks to you my stay in Germany has exceeded my expectations and the memories we created together will last forever in my mind.

Finally, I want to thank my boyfriend Andrea, my mum Nicole, and my sister Sara, for being always my greatest supporters and for showing me how proud they are. I love you so much.