Check for updates

Recessive variants in *ZNF142* cause a complex neurodevelopmental disorder with intellectual disability, speech impairment, seizures, and dystonia

Kamal Khan, MS [1,2,3], Michael Zech, MD [4,5], Angela T. Morgan, PhD [6],
David J. Amor, MBBS, PhD [6], Matej Skorvanek, MD, PhD [7,8], Tahir N. Khan, PhD[1,17],
Michael S. Hildebrand, PhD[6,9], Victoria E. Jackson, PhD [10], Thomas S. Scerri, MSc, DPhil [10],
Matthew Coleman, BA/BSc(Hons)[9], Kristin A. Rigbye, BSc(Hons)[9], Ingrid E. Scheffer, MBBS, PhD [9,11],
Melanie Bahlo, PhD [10], Matias Wagner, MD [4,5], Daniel D. Lam, PhD [4], Riccardo Berutti, PhD[12],
Petra Havránková, MD, PhD[13], Anna Fečíková, MD [13], Tim M. Strom, MD[5,12], Vladimir Han, MD[7,8],
Petra Dosekova, MD[7,8], Zuzana Gdovinova, MD[7,8], Franco Laccone, MD[14], Muhammad Jameel, PhD[2],
Marie R. Mooney, PhD [1], Shahid M. Baig, PhD [2,3], Robert Jech, MD, PhD[13], Erica E. Davis, PhD [1],
Nicholas Katsanis, PhD[1] and Juliane Winkelmann, MD[4,5,15,16]

Purpose: The purpose of this study was to expand the genetic architecture of neurodevelopmental disorders, and to characterize the clinical features of a novel cohort of affected individuals with variants in *ZNF142*, a C_2H_2 domain-containing transcription factor.

Methods: Four independent research centers used exome sequencing to elucidate the genetic basis of neurodevelopmental phenotypes in four unrelated families. Following bioinformatic filtering, query of control data sets, and secondary variant confirmation, we aggregated findings using an online data sharing platform. We performed in-depth clinical phenotyping in all affected individuals.

Results: We identified seven affected females in four pedigrees with likely pathogenic variants in *ZNF142* that segregate with recessive disease. Affected cases in three families harbor either nonsense or frameshifting likely pathogenic variants predicted to undergo nonsense mediated decay. One additional trio bears ultrarare missense variants in conserved regions of ZNF142 that are predicted to be damaging to protein function. We performed clinical comparisons across our cohort and noted consistent presence of intellectual disability and speech impairment, with variable manifestation of seizures, tremor, and dystonia.

Conclusion: Our aggregate data support a role for *ZNF142* in nervous system development and add to the emergent list of zinc finger proteins that contribute to neurocognitive disorders.

Genetics in Medicine (2019) 21:2532–2542; https://doi.org/10.1038/s41436-019-0523-0

Keywords: homozygosity mapping; developmental delay; ataxia; dolichocephaly; childhood apraxia of speech

INTRODUCTION

Zinc finger proteins (ZNF/zfp) comprise a diverse group of cellular effectors encoded by >1500 genes (HGNC database; https://www.genenames.org).[1,2] ZNF/zfp contain different arrays of zinc finger domains, among which the Cys_2His_2 (C_2H_2) motif is abundant in eukaryotic genomes, and is found in >700 proteins (Fig. S1).[1,2] The zinc finger proteins have established roles in the modulation of a wide range of cellular processes.[2] The central and peripheral nervous systems seem to be susceptible to ZNF/zfp dysfunction: an accumulating list of pathogenic variants in ZNF/zfps have been reported in

individuals affected by neurodevelopmental disorders, genetically and phenotypically diverse conditions impacting the central nervous system.[3] For example, recessive variants in *ZBTB11* cause a complex intellectual disability disorder accompanied by microcephaly, delayed motor milestones, facial hypotonia, and ataxia.[4] Furthermore, several ZNF/zfps have been proposed to be the causative genes in genomic disorders. For example, de novo variants in *ZBTB20* cause Primrose syndrome, offering genetic support of its involvement in the syndromic neurodevelopmental features of the 3q13.31 microdeletion syndrome.[5–8] Additionally, loss-of-

function (LOF) missense variants, deletions, partial duplications, and large chromosomal rearrangements at the *ZEB2* locus are associated with Mowat–Wilson syndrome, a condition characterized by distinctive facial features, intellectual disability, and congenital anomalies impacting multiple organ systems.[9–12]

Pathogenic variants in ZNF/zfp-encoding genes have also been reported in movement disorders, such as tremor and dystonia, clinical entities also considered to be neurodevelopmental disorders.[13] Notably, recent studies have identified dystonia as one component of a complex disorder, Gabriele–de Vries syndrome, caused by haploinsufficiency in Yin and yang 1 (*YY1*), a C_2H_2 zinc finger transcription factor.[14] However, variations in ZNF/zfps do not always impact cognitive function and can also cause isolated dystonia; for example, heterozygous missense variants as well as indels in *THAP1/DYT6*, a zinc finger protein with a conserved DNA-binding THAP domain, are associated frequently with dystonia in humans.[15,16] The mechanisms by which pleiotropy is governed in ZNF/zfp disorders remains poorly understood; however, transcriptomic and CHIP-Seq experiments offer clues about the spatiotemporal effects of ZNF/zfp dysfunction. As one step toward addressing this question, RNA-seq of striatum and cerebellum from a *Thap1* murine model showed dysregulation of pathways that are regulated in a tissue-dependent manner.[17] In aggregate, ZNF/zfp transcription factors likely induce broad phenotypic consequences in accordance with expansive downstream target genes involved in development and homeostasis.

Here, we report four unrelated families comprising seven affected females with an overlapping neurodevelopmental disorder hallmarked by intellectual disability and speech impairment, variably expressive seizures, tremor, and dystonia. Investigators in four research centers independently performed exome sequencing (ES) in these families with undiagnosed conditions. In all cases, we identified recessive variants in *ZNF142*, encoding zinc finger protein 142, as a causal gene associated with the pathology.

MATERIALS AND METHODS
Ethics statement, recruitment, and subject evaluation
Clinical phenotyping, ES, and molecular analyses were approved by the relevant institutional ethics committees from participating centers (Munich, Germany; Prague, Czech Republic; Kosice, Slovakia; Vienna, Austria; Faisalabad, Pakistan; Durham, NC; Melbourne, Australia). Signed informed consent for study procedures and publication of clinical and genetic findings was obtained from all the participants or their legal representatives. For seven affected individuals, extensive clinical evaluation including medical history interviews, neurological and physical examinations, and review of chart records was performed. We obtained saliva (using Oragene kits) or peripheral blood samples from affected individuals and their healthy family members following standard venipuncture protocols, and we extracted

DNA using standard phenol-chloroform methods or commercial kits: prepIT•L2P kit (DNA Genotek Inc, Ontario, Canada) for saliva and Qiagen QIAamp DNA Maxi Kit (Hilden, Germany) for blood. All participating research centers were connected through the public data sharing platform GeneMatcher (entry "ZNF142").[18]

Exome sequencing and variant analyses
Simplex, trio, quartet, or extended family-based exome sequencing (ES) (Fig. 1 and Fig. S2) was conducted using different commercial exome libraries and Illumina sequencing platforms; standard bioinformatics analyses are described in the Supplementary Data. Briefly, raw exome data were processed using standard read-mapping algorithms, aligner tools, quality control (QC) checks, and variant callers. Aligned reads were inspected with the Integrative Genomics Viewer (IGV, Broad Institute) and filtering of single-nucleotide and copy-number variants was conducted using objective criteria including variant type (e.g., nonsense, frameshift, missense), minor allele frequency (MAF) in relevant ethnicity-matched populations (e.g., gnomAD), in silico pathogenicity prediction tools (e.g., CADD, PolyPhen-2 for missense variants), and disease models (e.g., recessive, dominant, de novo).

Sanger sequencing
To verify candidate variants identified by ES and to test for segregation with disease, we performed bidirectional Sanger sequencing in all participating family members for whom DNA was available using BigDye Terminator chemistry on an ABI 3730 sequencer. Primer sequences are available upon request.

Homozygosity mapping (family C)
To identify homozygous regions shared between the two cases with ES data, we performed genome-wide homozygosity mapping. We generated a joint variant call file (vcf) of ES data using the GATK HaplotypeCaller and GenotypeGVCFs methods;[19] this data set consisted of the two cases, the unaffected mother, and two healthy siblings. We used HomozygosityMapper[20] to plot homozygosity stretches genome-wide using default parameters (Fig. 2a–c), and we used GeneDistiller 2014 and NCBI Genome Data Viewer to visualize the *ZNF142* locus (Fig. 2c, d).

RESULTS
Clinical reports
Family A
Subject A-II-2 (Fig. 1a; Table 1) was born at term after an uncomplicated pregnancy; she is the second child of healthy unrelated parents of Slovakian descent. Although birth parameters and neonatal adaptation were normal, the parents reported feeding difficulties. She first presented to the clinic during her second year of life for developmental delay concerns. She sat at 12 months of age and started walking at 24 months, albeit with balance problems. Her speech development was impaired; she spoke her first 2–3 word
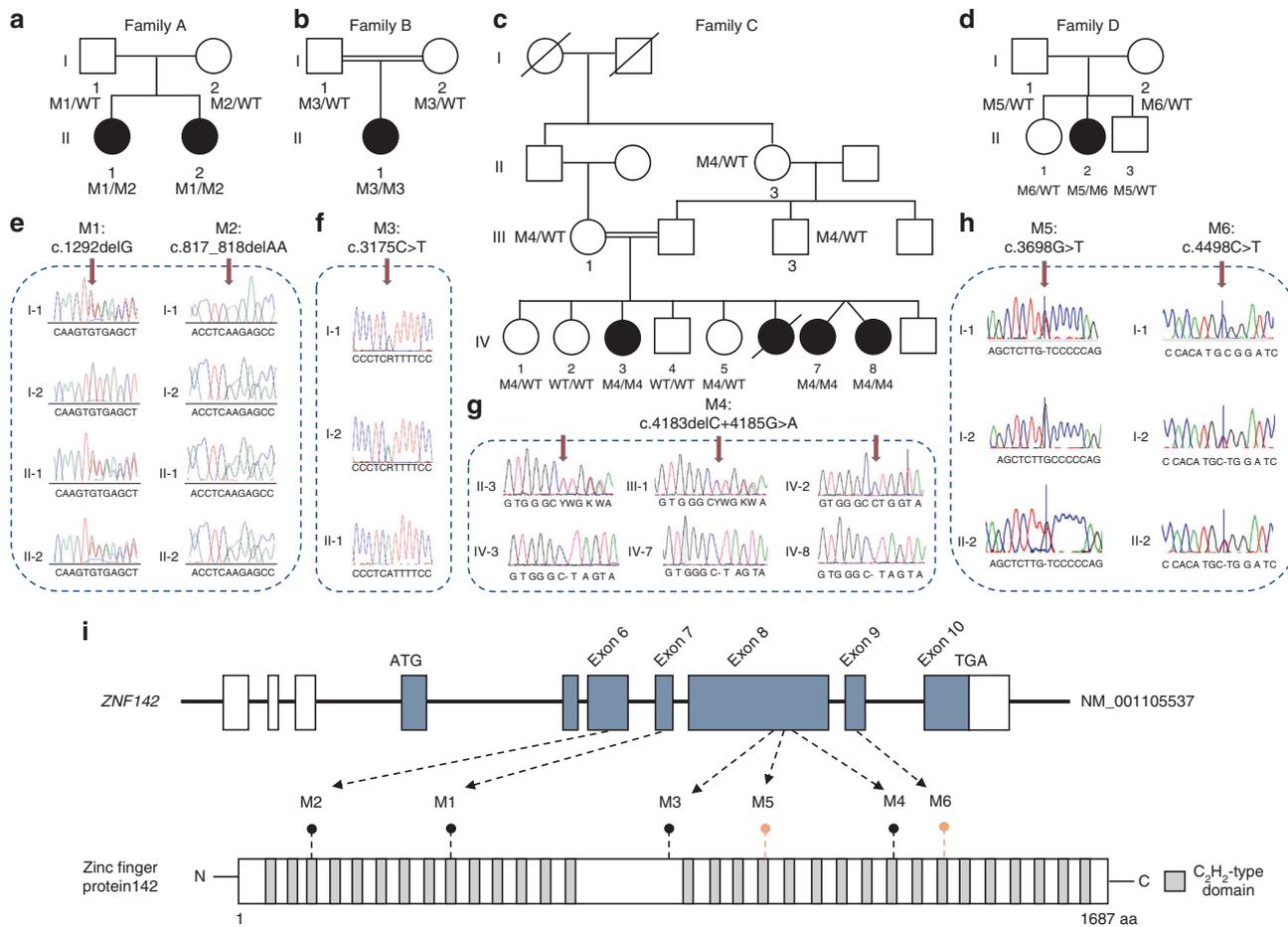
Fig. 1 Likely pathogenic variants in *ZNF142* are associated with syndromic neurodevelopmental phenotypes in four unrelated pedigrees. (a–d) Pedigrees of four families with syndromic neurodevelopmental symptoms segregating in an autosomal recessive pattern of inheritance. Double lines in pedigrees indicate consanguinity. Filled and unfilled circles/squares represent affected and unaffected individuals respectively, while circle/squares with diagonal lines indicate deceased individuals. Genotypes are represented as either WT/WT (wild type), MX/WT (heterozygous) or MX/MX (homozygous) for individuals with available genotypes. (e–h) Representative chromatograms are shown for each variant. Vertical arrows indicate variant position. (i) Top: schematic of the *ZNF142* locus with exons 6–9 harboring likely pathogenic variants. Boxes, exons; black line, introns; white, untranslated regions; blue shaded boxes, coding regions. Bottom: schematic representation of ZNF142 protein; gray rectangles represent predicted C_2H_2-type domains. Truncating alterations are indicated with black lollipops; missense variants are indicated with salmon-colored lollipops.

sentences at age 5. On at least one occasion, she suffered a generalized tonic–clonic seizure that ceased spontaneously. By the age of 4 years, she developed dystonic head posturing and tremulous movements of the arms. Over time, abnormal hyperkinetic movements, exaggerated during purposeful motor activities, became a prominent disabling symptom. Cognitive assessment at age 6 years revealed impaired mental development, requiring special schooling. Behavioral testing showed deficits in social domains and executive functions. Routine laboratory studies and additional examinations including cerebrospinal fluid analysis, electroencephalography, and tests for inborn errors of metabolism were nondiagnostic. Genetic testing ruled out Angelman syndrome, Rett syndrome, hereditary dystonia 1 and 6, and the most common ataxia disorders. Karyotyping and chromosomal microarray analyses were normal. A computed tomography scan of the head did not reveal structural anomalies. At her last clinical examination (age 33 years), subject A-II-2

displayed tremulous cervical dystonia, facial dyskinesia, action and postural tremor in her upper limbs, and an ataxic gait. She also had severe intellectual disability (IQ: 50) and profound speech problems with restricted vocabulary. Physical examination showed dolichocephaly.

Individual A-II-1 (Fig. 1a; Table 1) displayed a constellation of symptoms resembling that of her younger sister, A-II-2. Her medical history was marked by signs of global psychomotor deficits, involving fine and gross motor skill development (walking at 15 months with lower limb incoordination); language (single words spoken at 5 years); and limited socialization. She presented with hyperkinetic movement abnormalities and gait disturbances since early childhood. No seizures have been documented. Multiple biochemical assays and routine molecular testing all gave normal results. At her most recent clinical evaluation at 35 years of age, subject A-II-1 demonstrated segmental craniocervical dystonia, tremor of the head and upper limbs, gait
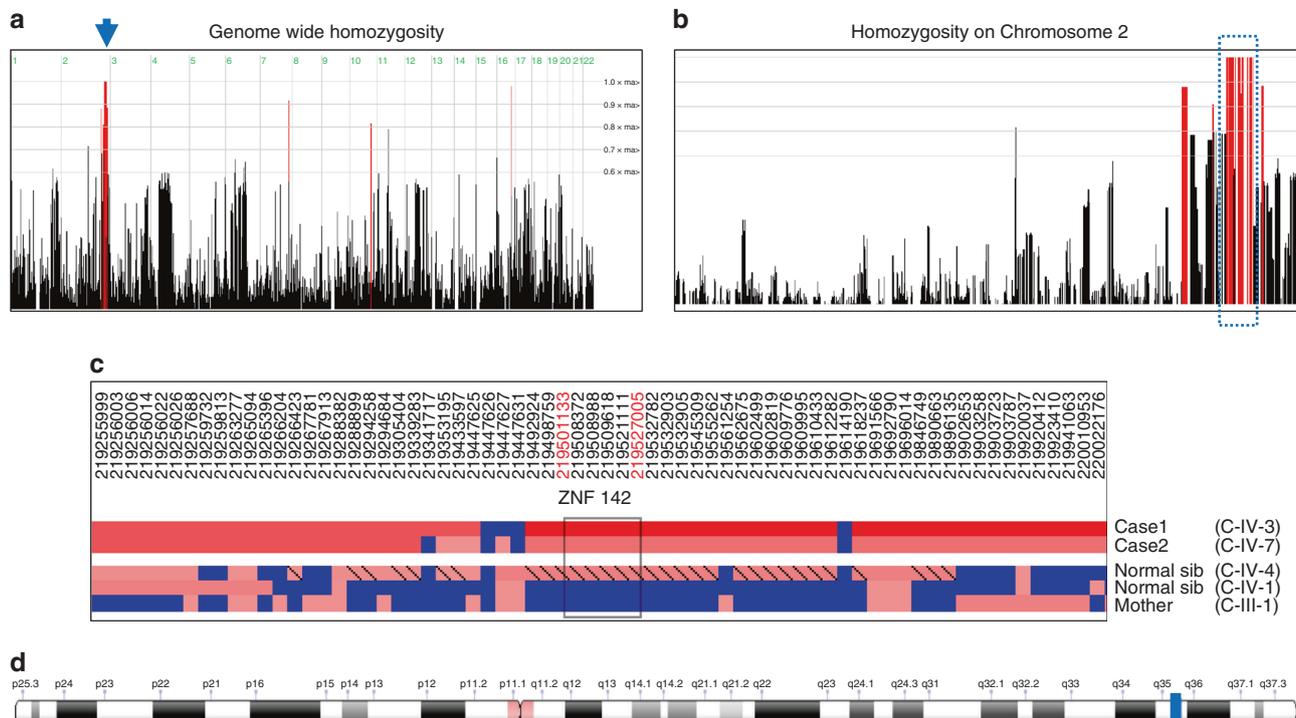
Fig. 2 Genome-wide homozygosity mapping in family C using exome sequencing data. (**a**) Stretches of homozygosity were mapped on jointly called exome sequencing (ES) data from five individuals (two affected, two unaffected siblings, and their unaffected mother) using HomozygosityMapper software. Black vertical lines, regions with <80% homozygosity exclusive to cases; red vertical lines, regions with >80% homozygosity exclusive to cases. The blue arrow indicates a stretch of homozygosity on chr2 harboring the *ZNF142* variant. (**b**) Enlarged view of chr2 indicating the homozygous genomic stretches shared between affected individuals. The blue dotted rectangle outlines the 10.6-Mb region of interest with a 100% homozygosity score. (**c**) Graphical representation of the homozygous region on chr2 generated with GeneDistiller. *ZNF142* is localized to the homozygous region shared between cases and differs from healthy siblings and their unaffected mother. Black rectangle and flanking hg19 coordinates (red, top) indicate the region containing *ZNF142*. Blue, heterozygous variant calls; red, homozygous regions with lighter red to darker red reflecting shorter to longer homozygous stretches, respectively. (**d**) Ideogram of chr2 generated by Genome Data Viewer showing the position of *ZNF142* (blue rectangle) at 2q35 (hg19: chr2: 219502640-219524355).

ataxia, severe intellectual disability (IQ: 52), speech impairment, and dolichocephaly.

Family B

Subject B-II-1 (Fig. 1b; Table 1) was the only child born to consanguineous Turkish parents following an uneventful full-term pregnancy. There were no reported perinatal problems and growth parameters were within normal limits. The postnatal interval was characterized by feeding difficulties and excessive crying. We noted a hyperkinetic movement disorder that became evident during the first year of life. She developed dystonic posturing of the neck, trunk, and all four limbs. We also saw a gradual progression of chorea-like movements of the arms. Additionally, all aspects of psychomotor development were delayed from early infancy. She achieved unsupported sitting, but she was never able to walk independently and was wheelchair bound. She was unable to speak, and she communicated using gestures. Although higher mental function was not formally assessed, her cognitive level was estimated to be moderately disabled. At the age of 2 years, she experienced an episode of tonic–clonic seizures. Electroencephalography reported excessive beta activity, with no clear evidence of epileptiform discharges.

Biochemical and metabolic markers were normal. Targeted genetic testing for mitochondrial disorders and hereditary dystonias did not yield any likely pathogenic variants. Brain magnetic resonance imaging (MRI) findings, which did not change appreciably over a five-year period (three to eight years of age), were nondiagnostic and consisted of nonspecific, subtle signal intensities in the lateral aspect of the left thalamus, subcortical gyrus pre- and postcentralis, subcortical occipital area, and left putamen. At last clinical examination (eight years old), we noted generalized dystonia combined with bilateral choreatic movements and minor ataxic features. There was still no expressive language, but she showed evidence of verbal comprehension. Craniofacial examination detected no specific dysmorphic signs.

Family C

We evaluated three affected females and their healthy mother from an extended multigenerational consanguineous Pakistani family (Fig. 1c; Fig. S2; Table 1). The proband (C-IV-3), presented with severe intellectual disability at her last clinical assessment (aged 18 years). She was born healthy subsequent to a full-term pregnancy with no known environmental contributing factors; she sat independently at the age of six

Table 1 Molecular and clinical characterization for seven individuals with biallelic variants in ZNF142

	Family A		Family B	Family C			Family D
	A-II-1	A-II-2	B-II-1	C-IV-3	C-IV-7	C-IV-8	D-II-2
Sex	Female	Female	Female	Female	Female	Female	Female
Age at most recent examination	35 years	33 years	8 years	18 years	13 years	13 years	10 years
Ethnicity	Slovakian	Slovakian	Turkish	Pakistani	Pakistani	Pakistani	European
Family history	Positive	Positive	Negative	Positive	Positive	Positive	Negative
Nucleotide alteration (NM_001105537.2; inheritance)	c.817_818delAA; c.1292delG; compound heterozygous	c.817_818delAA; c.1292delG; compound heterozygous	c.3175C>T; homozygous	c.4183delC; 4185G>A; homozygous	c.4183delC; 4185G>A; homozygous	c.4183delC; 4185G>A; homozygous	c.3698G>T; c.4498C>T compound heterozygous
Protein alteration (NP_001099007.1)	p.Lys273Glufs*32; p.Cys431Leufs*11	p.Lys273Glufs*32; p.Cys431Leufs*11	p.Arg1059*	p.Leu1395*	p.Leu1395*	p.Leu1395*	p.Cys1233Phe; p.Arg1500Trp
Neurodevelopment and function							
Developmental milestones	Sit: 10 months; walk: 15 months	Sit: 12 months; walk: 24 months	Sit: 4 years; wheelchair bound	Sit: 6 months; walk: 1 year	Sit: 15 months; walk:18 months	Sit: 15 months; walk: 18 months	Sit: 12 months; walk: 18 months
Cognition	Severely Impaired (IQ 52)	Severely impaired (IQ 50)	Moderately impaired; IQ not available	Severely impaired WIAS-R/FSIQ (59)	Severely impaired WISC-R/FISQ (40)	Severely impaired; not able to perform WISC-R	Low average verbal IQ (81), nonverbal IQ (82), and FSIQ (78)
Seizure(s) (Age at onset)	–	At least once, generalized tonic–clonic (1–2 years)	Once, generalized tonic–clonic (2 years)	7–8 times, tonic–clonic (4 months)	2x/week, tonic–clonic (13 days)	7–8 times, tonic–clonic (13 days)	–
Speech	Delayed: singular words at 5 years; last examination: cerebellar dysarthria with hypernasality, restricted vocabulary	Delayed: first 2–3 word sentences at 5 years; last examination: spastic speech with severely impaired speech initiation, restricted vocabulary	(No speech) no expressive language, verbal comprehension intact	Delayed speech; unable to speak in full sentences	Delayed speech; unable to speak in full sentences	Receptive and expressive speech delayed	Childhood apraxia of speech
Movement disorder characteristics and other neurological features							
Dystonia (age of onset)	Segmental (4 years)	Segmental (4 years)	Generalized (1 year)	–	–	–	–
Dystonia site of onset	Neck	Neck	Neck	NA	NA	NA	NA
Areas involved in dystonia (most recent examination)	Face, neck, upper limbs	Face, neck, upper limbs	Face, neck, trunk, all 4 limbs	NA	NA	NA	NA
Tremor	+	+	–	–	+	+	–
Ataxia	+	+	+	–	+	–	–

Table 1 continued

	Family A		Family B	Family C			Family D
	A-II-1	A-II-2	B-II-1	C-IV-3	C-IV-7	C-IV-8	D-II-2
Other neurological features (age of onset)	–	–	Trunk hypotonia, choreatic movements	–	–	–	Increased tone and reflexes in lower limbs (3 years)
Brain MRI	Not performed	Not performed	Signal intensities	Normal	Not performed	Not performed	Mild thinning of the posterior aspect of the body of the corpus callosum with associated mild reduction in white matter volume
Other features	Dolichocephaly	Dolichocephaly	–	Dolichocephaly	–	Dolichocephaly	–

FSIQ full scale intelligence quotient, *IQ* intelligence quotient, *MRI* magnetic resonance image, *NA* not applicable, *WIAS-R* Wechsler Intelligence Scale for Adults—Revised, *WISC-R* Wechsler Intelligence Scale for Children—Revised, (+) yes, (−) no.

months, and started walking at the age of one year. At the age of four months, she displayed her first seizures, which reoccurred seven to eight times during her first three years of life; these have since abated without any medical interventions. During her seizure episodes, she became unconscious for 10–20 minutes. At her last clinical assessment, she also presented with severe speech impairment. She displays dolichocephaly with a head circumference of 53 cm (within normal range for age, 51–57 cm).[21] Her neurological assessment showed no symptoms of dystonia, ataxia, or tremors; MRI imaging of the brain showed no significant structural abnormalities.

The affected dizygotic twin female siblings in family C had similar features to the index case at last clinical assessment (aged 13 years). They presented with severe cognitive impairment and tonic–clonic seizures starting in the neonatal period (age 13 days). They were born healthy after a full-term pregnancy with no known teratogenic exposures, but shortly thereafter displayed developmental delay. They sat independently at the age of 15 months and started walking at 18 months. Individual C-IV-7 experiences seizures at variable frequency, ranging from twice a week to five to seven times per day with unknown triggering factors. The use of anticonvulsants stabilizes her condition during seizure episodes; however, the medications show no effect on the frequency of seizures. However, individual C-IV-8 displayed a total of seven to eight incidences of seizures in her lifetime. During these episodes, both affected twins became unconscious for 10–20 minutes. They also presented severe speech impairment and are unable to speak in full sentences. At their last clinical assessment, head circumference for both twins was within the normal range (C-IV-7, 53 cm; C-IV-8, 54 cm; normal, 49–55 cm),[21] but C-IV-8 displayed dolichocephaly. Neurological assessment revealed ataxia and generalized tremors in C-IV-7, whereas C-IV-8 displayed tremors and no other neurological features; MRI was not performed for either individual.

Their mother (C-III-I) had normal cognitive function, reported that she met all her developmental milestones, and had not experienced seizures or any other neurological symptoms. The three affected females have five healthy siblings, and a sixth deceased sibling who experienced long-lasting seizure episodes ranging from a few minutes to a few hours and was deceased at the age of nine years after experiencing an overnight seizure attack (Fig. 1c; Fig. S2).

Family D

Subject D-II-2 is a 10-year-old girl, the second of three children born to healthy parents (Fig. 1d; Table 1). Her older sister and younger brother are healthy. She was born at term by forceps delivery after an uncomplicated pregnancy and there were no neonatal complications. Her early development was mildly delayed: she sat at 12 months and walked at 18 months and was slow to acquire language. At age 3 years she was diagnosed with delayed motor development and showed brisk reflexes in the lower limbs and tightening of the

calves, suggestive of mild spastic diplegia. The main developmental concern reported by the parents has been a severe and persistent speech disorder, namely childhood apraxia of speech, which was severe from ages 3 to 8 years and has only begun to resolve in the past two years. She attends regular school but neuropsychological assessment showed low average IQ (full scale IQ [FSIQ] 78). She has not had any major physical health problems, has not had seizures, and has normal hearing and vision. Height, weight, and head circumference are all at the 90th centile. On examination she was nondysmorphic and could walk and run. She had normal muscle power but was hyperreflexive in the lower limbs with sustained clonus and increased tone. Upper limb reflexes are relatively brisk. An MRI showed mild thinning of the posterior aspect of the body of the corpus callosum with associated mild reduction in white matter volume. Chromosome microarray was normal.

Genomic studies identify rare biallelic variants in ZNF142

To investigate the genetic basis of disease in family A, we performed ES in a quartet paradigm. We achieved an average target read depth ranging from 114× to 149× across the four individuals, with 98.2% to 98.9% of coding regions covered by ≥20 reads (Table S1). Subsequent to our bioinformatic filtering criteria, we identified a single candidate gene, *ZNF142*, bearing likely pathogenic compound heterozygous variants that segregated with the phenotype. Both affected siblings (A-II-1 and A-II-2, Fig. 1a; Table 1; Table S2) harbored variants predicted to cause a frameshift; these included a 2-bp deletion (NM_001105537.2: c.817_818delAA; p.Lys273Glufs*32) and a 1-bp deletion (c.1292delG; p. Cys431Leufs*11; Fig. 1a, e, i) predicted to undergo nonsense mediated decay. We confirmed these results using bidirectional Sanger sequencing. Neither variant was reported in public databases (>140,000 individuals in ExAC and gnomAD v2.1) or our in-house control ES data collections (10,000 individuals).

We next considered the known biological function of *ZNF142*. It is a predicted transcription factor expressed ubiquitously across adult tissues, with the highest expression reported in the brain, specifically the cerebellum (GTEx Portal; gtexportal.org). This gene has not been implicated previously in human genetic disorders, but a mouse mutant with homozygous ablation of the orthologous locus (*Zfp142*) has behavioral and neurological phenotypes (Mouse Genome Informatics and the International Mouse Phenotyping Consortium [IMPC]; MGI ref.ID: J:211773).

Next, we queried *ZNF142* in our in-house exome database (*n* = 10,000 individuals with various neurological and non-neurological phenotypes of diverse ethnicities; Munich). We identified an individual affected with neurodevelopmental deficits who had no previously established genetic diagnosis (family B; B-II-1; Fig. 1b; Table 1). She had a homozygous nonsense variant (c.3175C>T; p.Arg1059*) predicted to truncate the latter half of the polypeptide (Fig. 1b, f, i; Table 1; Table S2). We validated this finding with Sanger

sequencing in the trio; the variant segregated in an autosomal recessive fashion from heterozygous carrier parents. The nonsense change, albeit rare, was documented in the NCBI database (rs546151500) and was also reported in two heterozygous individuals in gnomAD (MAF: 8.022×10^{-6}).

Reanalysis of the family B proband's ES data identified an additional rare homozygous variant in a known human disease gene, *SLC19A3* (NM_025243.3: c.952G>A; p. Ala318Thr; Table S2), disruption of which causes autosomal recessive thiamine metabolism dysfunction syndrome 2 (THMD2; MIM 607483).[22] THMD2 is hallmarked by recurrent bouts of subacute encephalopathy, often progressing to coma and death in early childhood.[22] The *SLC19A3* change, which we confirmed by Sanger sequencing, was located in a contiguous homozygous stretch that also contained the *ZNF142* variant. Further, the p.Ala318Thr variant was absent from gnomAD v2.1, and absent from morbid human gene databases, and thus categorized as a variant of uncertain significance (VUS). To investigate a possible contribution of the *SLC19A3* change to phenotype, we performed repeated brain MRI evaluation, but observed no degenerative processes or signal alterations as they are found typically in individuals with THMD2. Moreover, empirical treatment of B-II-1 with high doses of biotin/thiamine, which is known to ameliorate cases with LOF variants in *SLC19A3*,[22,23] resulted in no clinical improvement thereby reducing further the likelihood of this locus as a major phenotype contributor. We identified pathogenic heterozygous variants in two different known disease-causing genes (*GMMPB* and *MUT*), but both loci cause non–phenotype overlapping recessive conditions (Table S3).

To identify molecular lesions that might contribute to the pathology of family C, we conducted ES on six individuals (Fig. 1c; Fig. S2). We obtained a mean target coverage of 67–91× with 94% of bases covered ≥20× for all individuals (Table S1). Based on the recurrence of similar phenotypes (four of nine siblings; Fig. 1c; Table 1) and unaffected consanguineous parents, we posited an autosomal recessive mode of inheritance. First, we performed homozygosity mapping with ES data. We recalled variants among the five individuals from the same branch of the pedigree jointly (C-III-1, C-IV-1, C-IV-3, C-IV-4, and C-IV-7), and identified homozygous genomic regions shared between affected individuals. As expected, all four siblings displayed homozygosity consistent with previous reports for consanguineous pedigrees (9–14%) (ref. [24]). We identified four homozygous stretches localized to chromosomes 2, 7, 10, and 16 with >80% of the maximal homozygosity score (Fig. 2a). Among these regions, the segment on chr2q was the only region to achieve the maximal homozygosity score (Fig. 2a–d). This region contains three blocks encompassing 10.6 Mb (chr2:214012405-224620246; hg19), contains 340 homozygous markers, and includes 84 protein-encoding genes.

Independently of the homozygosity analysis, we prioritized rare (MAF < 1%) functional variants that were either homozygous or compound heterozygous and shared between

affected individuals. There were no loci with compound heterozygous variants fulfilling these criteria. However, we found rare homozygous changes present in two genes, *RUFY4* and *ZNF142*, both present within the region on 2q35 identified by homozygosity mapping. First, we considered the missense variant in RUN and FYVE domain containing 4 (*RUFY4*; NM_198483.3: p.Asn194Ile; Table S2). Although this variant is present at MAF 0.01% in gnomAD (320/280,382 alleles), it is enriched in the South Asian population (MAF = 1.03%, including three individuals in homozygosity). Due to the high frequency of this variant in the South Asian population, we considered it unlikely to cause the pathology.

Next, we turned our attention to the homozygous likely pathogenic variant in *ZNF142*, encoding p.Leu1395*, which is the combination of a deletion and a substitution (c.4183delC +c.4185G>A; Table S2). The deletion variant (c.4183delC) results in substitution of leucine to tryptophan and subsequent frameshift (p.Leu1395Trpfs*2) and the substitution (c.4185G>A) is a synonymous variant. Neither change is present in homozygosity in gnomAD, and each variant is found in heterozygosity once in >249,000 alleles. Segregation analysis in 17 available family members from the pedigree demonstrate that these two changes are in *cis* (Fig. 1c, g; Fig. S2). The variant is predicted to truncate 292 amino acids from the C-terminal encoding region of the putative protein (Fig. 1i), removing 9 of 31 DNA-binding C_2H_2-type domains, and is likely pathogenic.

To uncover the genetic cause of pathology in family D, we filtered variants from ES data (Table S1) of the trio (unaffected parents D-I-1 and D-I-2, and affected proband II-2). We filtered ES data for de novo variants, and no likely deleterious changes met our filtering criteria (Table S3). Next, analysis of the ES data following a recessive paradigm did not identify homozygous variants, however, we identified compound heterozygous variants in *ZNF142* (Table S2): c.3698G>T (p.Cys1233Phe) inherited from the father (D-I-1) and c.4498C>T (p.Arg1500Trp) inherited from the mother (D-I-2) (Fig. 1d, h). Both variants are present in public databases in heterozygosity (1 of 249,582 and 4 of 280,716 alleles in gnomAD, respectively), but homozygotes have not been reported for either change. Both variants reside in C_2H_2-type zinc finger domains (Fig. 1i) and are predicted to impact the function of ZNF142 (PolyPhen-2 scores: 0.998 and 0.999; SIFT scores: 0 and 0.005; and CADD scores: 31 and 26, respectively). Sanger analysis showed that variants segregate with disease (Fig. 1d, h).

DISCUSSION

Here, we report four unrelated families with seven affected females who exhibit overlapping clinical features and harbor likely pathogenic recessive variants in *ZNF142*. We conducted investigations that include (1) recruitment and clinical evaluation of patients with neurodevelopmental disorders and their family members, (2) pedigree-based ES and bioinformatic filtering, (3) genome-wide homozygosity

mapping (family C, Fig. 2a–d), and (4) community-wide data sharing. The LOF variants identified in families A, B, and C are predicted to result in premature stop codons; we expect the resulting products to be targeted for degradation through nonsense mediated decay[25,26] or, less likely, to lead to the generation of truncated proteins that lack functional C_2H_2 domains. The missense variants identified in family D localize to the conserved regions of ZNF142 and are predicted to be pathogenic (Fig. S3; Table S2).

The families described in this study were assessed as part of research projects with differing phenotypic foci (families A and B, movement disorders; family C, syndromic intellectual disability; family D, childhood apraxia of speech). However, reverse phenotyping revealed that all affected individuals had a strikingly similar course of their illness. Unifying symptoms of the disorder include moderate to severe cognitive impairment (7/7), speech deficits (7/7), motor impairment (7/7), variably penetrant tonic–clonic seizures (5/7), tremor (4/7), and dystonia (3/7). The complex phenotypic traits can potentially be explained by the widespread roles of ZFN/zfps, which impact the regulation of diverse neuronal pathways and function;[1,4,17,27] we do not know whether observed variable expressivity in our patients is the product of genetic background differences or stochastic factors. Additional families with pathogenic variants at this locus will be required to understand the range of ZNF142-related phenotypes. We also note that the family B proband had a more severe presentation in terms of movement disorder and motor incoordination in comparison with other cases. She harbors a homozygous VUS in *SLC19A3* in addition to the *ZNF142* variant. We speculate that an additive or epistatic effect between the two loci might contribute to the observed clinical severity, although there are currently no known interactions between the two genes that support this possibility.

Our work expands the functional role of ZNF/zfps in neurodevelopment, however, the precise mechanisms through which ZNF142 impacts neuronal function are unclear. Although ZNF142 is part of the largest subgroup of ZNF/zfps (Fig. S1), it does not have a recent common structure with most other C_2H_2 family members. To investigate the relationship of ZNF142 to other genes that cause neurodevelopmental disorders, we performed a multiple sequence alignment of 492 of the 720 C_2H_2 proteins in the HGNC database that had UniProt/SwissProt annotations using COBALT[28] (Fig. 3a). We examined protein similarities based on (1) the 492 annotated family members and (2) the ZNF142 adjacent branches. We found that of the 492 annotated C_2H_2 subgroup members, 58 (18%) are linked to clinical synopses (Table S4; OMIM; https://www.omim.org/). Additionally, 67% of OMIM-associated genes are involved in neurodevelopmental disorders. Notably, three of the four closest subfamily members to ZNF142 have been implicated in syndromic intellectual disability (Fig. 3b; Table S4).[29–31] For example, recessive missense variants in *ZNF407* result in cognitive impairment, hypotonia, dysmorphic features, and digit defects.[30]
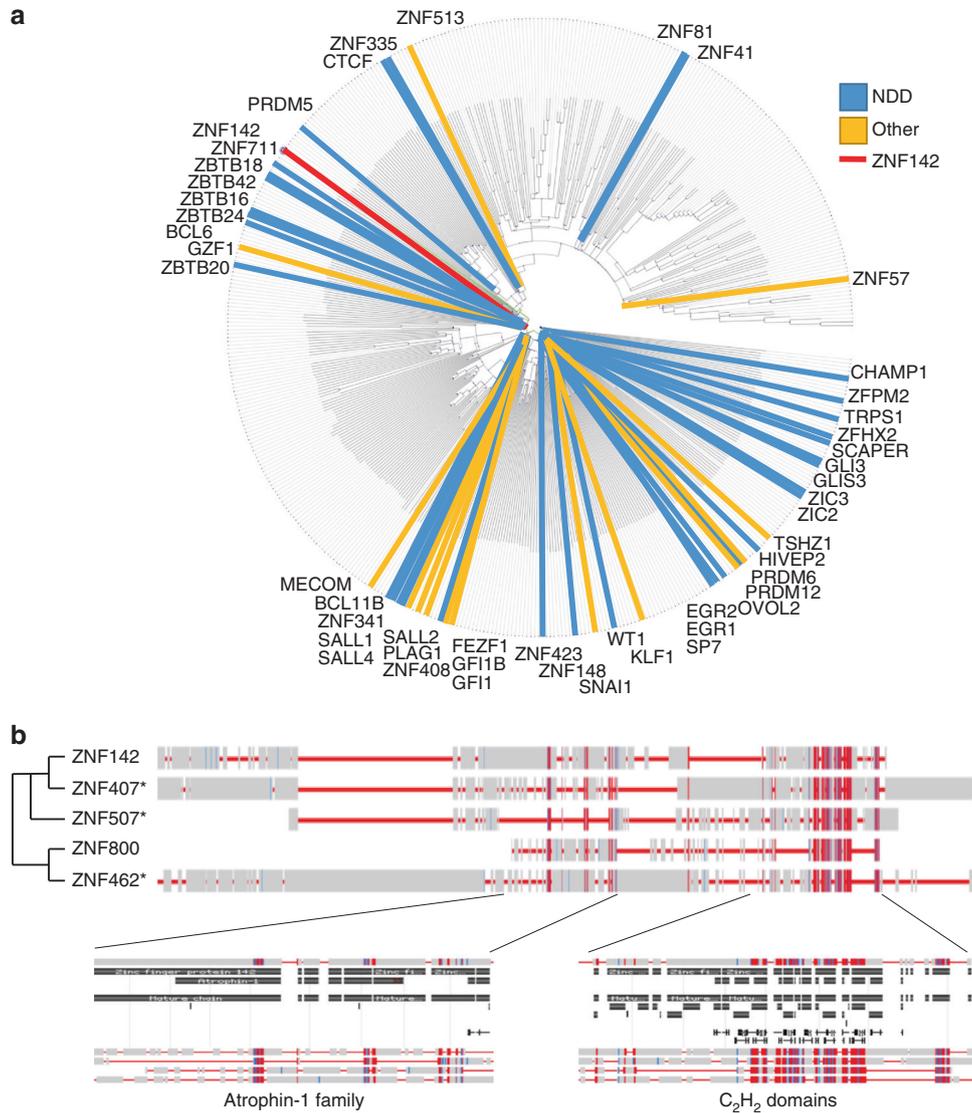
Fig. 3 Protein similarity between ZNF142 and other C$_2$H$_2$ family members. Seven hundred twenty C$_2$H$_2$ domain-containing members of zinc finger proteins queried from the HUGO gene nomenclature committee were mapped to protein accession numbers with BioMart. Mapped proteins were used as input for a conserved domain and local sequence similarity multiple protein alignment using COBALT (maximum sequence difference set to 0.85 and using the Grishin distance metric). (**a**) Four hundred ninety-two mapped proteins are displayed as a COBALT distance circular map sorted by their structural distance. In addition to ZNF142, proteins with OMIM clinical synopses are marked on the map, along with their designation as involved in neurodevelopmental processes (NDD; blue) or other processes (yellow; see Table S4). (**b**) ZNF142 is aligned with the proteins in the adjacent branches of the cladogram and conserved protein domains for the Atrophin-1 family structure and C$_2$H$_2$ domain are shown. Asterisk denotes proteins without annotations in OMIM, but with support in peer-reviewed literature for an NDD designation.

It is conceivable that lack of ZNF142 may alter spatiotemporally controlled transcriptional programs. This has been shown for THAP1, wherein genes involved in eIF2α signaling, neuron projection development, axonal guidance signaling, synaptic long-term depression, and mitochondrial dysfunction were reported to be dysregulated significantly.[17] Future studies on downstream target genes in patient cell lines or in vivo systems will shed light on roles played by ZNF142 and promote an understanding of the molecular pathogenesis of disease. Although we recognize that a limitation of our study is the small cohort size, we predict that the condition is underidentified in clinical practice. *ZNF142* molecular testing in individuals manifesting with a phenotypic spectrum overlapping speech impairment, epilepsy, and movement disorder will likely improve the ability to perform improved clinical and molecular studies.

SUPPLEMENTARY INFORMATION

The online version of this article (https://doi.org/10.1038/s41436-019-0523-0) contains supplementary material, which is available to authorized users.

DISCLOSURE

N.K. is a paid consultant for and holds significant stock in Rescindo Therapeutics, Inc. The other authors declare no conflicts of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Cassandri M, Smirnov A, Novelli F, et al. Zinc-finger proteins in health and disease. Cell Death Discov. 2017;3:17071.
2. Laity JH, Lee BM, Wright PE. Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol. 2001;11:39–46.
3. Sun Y, Hu D, Liang J, et al. Association between variants of zinc finger genes and psychiatric disorders: systematic review and meta-analysis. Schizophr Res. 2015;162:124–137.
4. Fattahi Z, Sheikh TI, Musante L, et al. Biallelic missense variants in ZBTB11 can cause intellectual disability in humans. Hum Mol Genet. 2018;27:3177–3188.
5. Cordeddu V, Redeker B, Stellacci E, et al. Mutations in ZBTB20 cause Primrose syndrome. Nat Genet. 2014;46:815–817.
6. Mattioli F, Piton A, Gerard B, et al. Novel de novo mutations in ZBTB20 in Primrose syndrome with congenital hypothyroidism. Am J Med Genet A. 2016;170:1626–1629.
7. Rasmussen MB, Nielsen JV, Lourenco CM, et al. Neurodevelopmental disorders associated with dosage imbalance of ZBTB20 correlate with the morbidity spectrum of ZBTB20 candidate target genes. J Med Genet. 2014;51:605–613.
8. Herve B, Fauvert D, Dard R, et al. The emerging microduplication 3q13.31: expanding the genotype-phenotype correlations of the reciprocal microdeletion 3q13.31 syndrome. Eur J Med Genet. 2016;59: 463–469.
9. Baxter AL, Vivian JL, Hagelstrom RT, et al. A novel partial duplication of ZEB2 and review of ZEB2 involvement in Mowat-Wilson syndrome. Mol Syndromol. 2017;8:211–218.
10. Yuan H, Zhang L, Chen M, et al. A de novo triplication on 2q22.3 including the entire ZEB2 gene associated with global developmental delay, multiple congenital anomalies and behavioral abnormalities. Mol Cytogenet. 2015;8:99.
11. Buraniqi E, Moodley M. ZEB2 gene mutation and duplication of 22q11.23 in Mowat-Wilson syndrome. J Child Neurol. 2015;30:32–36.
12. Ghoumid J, Drevillon L, Alavi-Naini SM, et al. ZEB2 zinc-finger missense mutations lead to hypomorphic alleles and a mild Mowat-Wilson syndrome. Hum Mol Genet. 2013;22:2652–2661.
13. Domingo A, Erro R, Lohmann K. Novel dystonia genes: clues on disease mechanisms and the complexities of high-throughput sequencing. Mov Disord. 2016;31:471–477.
14. Gabriele M, Vulto-van Silfhout AT, Germain PL, et al. YY1 haploinsufficiency causes an intellectual disability syndrome featuring transcriptional and chromatin dysfunction. Am J Hum Genet. 2017;100: 907–925.
15. Fuchs T, Gavarini S, Saunders-Pullman R, et al. Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia. Nat Genet. 2009;41:286–288.
16. Xiromerisiou G, Houlden H, Scarmeas N, et al. THAP1 mutations and dystonia phenotypes: genotype phenotype correlations. Mov Disord. 2012;27:1290–1294.
17. Zakirova Z, Fanutza T, Bonet J, et al. Mutations in THAP1/DYT6 reveal that diverse dystonia genes disrupt similar neuronal pathways and functions. PLoS Genet. 2018;14:e1007169.
18. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36:928–930.
19. 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. 2013;43:11 10 11–33.
20. Seelow D, Schuelke M, Hildebrandt F, Nurnberg P. HomozygosityMapper —an interactive approach to homozygosity mapping. Nucleic Acids Res. 2009;37:W593–599.
21. Nellhaus G. Head circumference from birth to eighteen years. Practical composite international and interracial graphs. Pediatrics. 1968;41: 106–114.
22. Zeng WQ, Al-Yamani E, Acierno JS Jr., et al. Biotin-responsive basal ganglia disease maps to 2q36.3 and is due to mutations in SLC19A3. Am J Hum Genet. 2005;77:16–26.
23. Debs R, Depienne C, Rastetter A, et al. Biotin-responsive basal ganglia disease in ethnic Europeans with novel SLC19A3 mutations. Arch Neurol. 2010;67:126–130.
24. Woods CG, Cox J, Springell K, et al. Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. Am J Hum Genet. 2006;78:889–896.
25. Kervestin S, Jacobson A. NMD: a multifaceted response to premature translational termination. Nat Rev Mol Cell Biol. 2012;13:700–712.
26. Lindeboom RG, Supek F, Lehner B. The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat Genet. 2016;48: 1112–1118.
27. Bradshaw NJ, Porteous DJ. DISC1-binding proteins in neural development, signalling and schizophrenia. Neuropharmacology. 2012;62:1230–1241.
28. Papadopoulos JS, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics. 2007;23: 1073–1079.
29. Hempel M, Cremer K, Ockeloen CW, et al. De novo mutations in CHAMP1 cause intellectual disability with severe speech impairment. Am J Hum Genet. 2015;97:493–500.
30. Kambouris M, Maroun RC, Ben-Omran T, et al. Mutations in zinc finger 407 [ZNF407] cause a unique autosomal recessive cognitive impairment syndrome. Orphanet J Rare Dis. 2014;9:80.
31. Talkowski ME, Rosenfeld JA, Blumenthal I, et al. Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. Cell. 2012;149:525–537.

[1]Center for Human Disease Modeling, Duke University Medical Center, Durham, NC, USA. [2]Human Molecular Genetics Laboratory, Health Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan. [3]Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan. [4]Institute of Neurogenomics, Helmholtz Zentrum München, Munich, Germany. [5]Institut für Humangenetik, Technische Universität München, Munich, Germany. [6]Murdoch Children's Research Institute and University of Melbourne Department of Paediatrics, Royal Children's Hospital, Parkville, Australia. [7]Department of Neurology, P.J. Safarik University, Kosice, Slovak Republic. [8]Department of Neurology, University Hospital of L. Pasteur, Kosice, Slovak Republic. [9]Epilepsy Research Centre, Department of Medicine, University of Melbourne, Austin Health, Heidelberg, VIC, Australia. [10]Population Health and Immunity Division, The Walter and Eliza Hall Institute of Medical Research, and University of Melbourne Department of Medical Biology and School of Mathematics and Statistics, Parkville, VIC, Australia. [11]University of Melbourne Department of Paediatrics, Royal Children's Hospital, and Florey and Murdoch Children's Research Institute, Parkville, VIC, Australia. [12]Institut für Humangenetik, Helmholtz Zentrum München, Munich, Germany. [13]Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine, Charles University and General Faculty Hospital, Prague, Czech Republic. [14]Institute of Medical Genetics, Medical School of Vienna, Vienna, Austria. [15]Lehrstuhl für Neurogenetik, Technische Universität München, Munich, Germany. [16]Munich Cluster for Systems Neurology, SyNergy, Munich, Germany. [17]Present address: Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Pakistan