The American Journal of Human Genetics, Volume *109*

Supplemental information

TCEAL1 loss-of-function results in an X-linked dominant neurodevelopmental syndrome and drives the neurological disease trait in Xq22.2 deletions

Hadia Hijazi, Linda M. Reis, Davut Pehlivan, Jonathan A. Bernstein, Michael Muriello, Erin Syverson, Devon Bonner, Mehrdad A. Estiar, Ziv Gan-Or, Guy A. Rouleau, Ekaterina Lyulcheva, Lynn Greenhalgh, Marine Tessarech, Estelle Colin, Agnès Guichet, Dominique Bonneau, R.H. van Jaarsveld, A.M.A. Lachmeijer, Lyse Ruaud, Jonathan Levy, Anne-Claude Tabet, Rafal Ploski, Małgorzata Rydzanicz, Łukasz Kępczyński, Katarzyna Połatyńska, Yidan Li, Jawid M. Fatih, Dana Marafi, Jill A. Rosenfeld, Zeynep Coban-Akdemir, Weimin Bi, Richard A. Gibbs, Grace M. Hobson, Jill V. Hunter, Claudia M.B. Carvalho, Jennifer E. Posey, Elena V. Semina, and James R. Lupski

SUPPLEMENTAL NOTE

SUPPLEMENTAL NOTE: CASE REPORTS

Individual 1 was a 5-year-old male with hypotonia, especially of the lower limbs, unsteady gait, and global developmental delay with particular difficulties in expressive language. He was born at 37 weeks as the first child to healthy parents via C-section due to maternal HELLP (Hemolysis, Elevated Liver enzymes, and Low Platelets) syndrome and breech presentation. There were no prenatal exposures to medications or drugs reported. He walked unassisted at 18 months of age; his gait was broad based with small steps and toe-walking and asymmetric running characterized by dragging of the right foot. At 4 years chronological age, fine motor, gross motor, and receptive language skills were in the 20-24 month level while expressive language was at a 15 month level. He had a handful of single words but his speech was primarily babbling and he used signs and a communication tablet. Some autistic behavior including repetitive/self-stimulating behavior and episodes of frustration were noted, but he was very social with good eye contact. Ophthalmology exam identified intermittent high frequency small amplitude nystagmus with a large latent component and large increase in amplitude and frequency in lateral gazes, intermittent exotropia, and mild myopia with astigmatism and the fundus exam identified a crowded disc and tortuous vessels. He had recurrent otitis media but subsequently had normal hearing. Height was normal at 25-50th centile and weight was 90-97th centile. Physical exam identified mildly dysmorphic facial features with a widow's peak, broad forehead, deep set eyes, mildly coarse facial features, and mildly low-set ears (Figure 1A, B), and a single café-au-lait macule on the upper right leg. Brain MRI at 34 months of age was essentially normal with possible mild foreshortening of the corpus callosum (CC). The CC thickness, myelination of the internal capsule, optic nerves, and vestibulo-cochlear nerves were appropriate for age (Figure 2A). Spine MRI at the same age (34 months) showed mild scoliosis, which may have been positional or secondary to hypotonia, and a prominent central canal. Metabolic workup including urine organic acids, plasma amino

acids, acylcarnitine, very long chain fatty acids, and serum transferrin isoelectric focusing, as well as chromosomal microarray and Fragile X testing were normal.

Trio analysis of ES data identified a *de novo* hemizygous nonsense variant in *TCEAL1*, c.447G>A p.(Trp149*), which was not present in gnomAD. Sanger sequencing confirmed the hemizygous variant in the proband and its absence in both parents. Based on its location within the Xq22 critical region, robust expression in brain, and absence of loss-of-function variants within gnomAD, this was considered a good candidate to potentially explain the observed clinical phenotype in individual 1.

Individual 2 was a 10-year-old male with generalized hypotonia and severe global developmental delay. He was born at 41 weeks via C-section after a failed induction; pregnancy was complicated by maternal PUPPP (Pruritic Urticarial Papules and Plaques of Pregnancy). He sat unaided at the age of 2 years and 3 months, but never crawled or walked. He demonstrated some developmental regression, having had a handful of single words as a young child, but was later non-verbal. He displayed an array of unusual, repetitive behaviors from an early age. These included abnormal hand movements, self-stimulating behaviors (some suggestive of gratification behavior), breath-holding episodes, hyperventilation episodes, and severe aerophagia (swallowing air). He also had a disturbed sleep cycle with sleep deprivation and more recently developed self-harming behaviors. He developed a seizure disorder in middle childhood. Ophthalmological examination identified hyperopia and astigmatism. He had longstanding problems with gut dysmotility and constipation, ultimately resulting in an ileostomy and colostomy, and a small epigastric hernia. He also had recurrent lung/airway infections that were found to be secondary to poor swallow and micro-aspirations necessitating gastric tube insertion, recurrent thumb dislocation, and episodic hyponatremia. Physical exam at 3 years of age found height at 25-50th centile, weight >99th centile, and head circumference at the 91-98th centile as well as mildly dysmorphic features including long palpebral fissures, brachycephaly, prominent bow-shaped upper lip, tapered fingers and shortened thumb (Figure 1C-D). Brain MRI at 10 years of age was normal.

Trio analysis identified a *de novo* hemizygous indel variant, c.299_302del p.(Gly100Alafs*22) (Figure 3), which was not present in gnomAD. Sanger sequencing was not able to be performed but the variant was independently identified as a high quality hemizygous *de novo* variant by two separate studies (DDD and 100KGP projects) in both exome and genome sequencing.

Individual 3 was a 17-year-old male presenting with global developmental delay, hypotonia, unsteady gait and growth retardation. He is the second child of a healthy unrelated couple. The pregnancy was uneventful and he was born at term, with a weight at 3380g (53rd centile), a length at 51cm (72nd centile) and a head circumference at 35.5cm (79th centile) He showed a global developmental delay in the first year, with a sitting at 15 months and global hypotonia. At evaluation, he could stand and walk some steps with assistance, but with substantial ataxia. He had no language but could communicate by eye contact and by gestures of intention. He had never had seizures. He had a history of recurrent otitis media but with normal hearing. He had some autistic traits with repetitive and stereotyped behaviors and was treated by Risperidone, but he was social and didn't have a major behavioral disorder. Ophthalmology exam identified strabismus and astigmatism. He had difficulty chewing and constipation. His height was between -1 and -2SD until 5 y/o when he showed a statural shift. The endocrine investigations showed normal growth hormone secretion but low levels of IGF1. At 17 years-old, he has short stature (145cm, -4SD) with normal weight and head circumference (55.5cm). He also had small testicles, with no pubertal onset. Physical examination identified dysmorphic facial features with deep set eyes, mildly coarse facial features, microdontia and sparse teeth. He also had very bright blue eyes, which contrasted with the dark eyes of both of his parents. A complete metabolic workup, including urine organic acids, and plasma and urine amino acids, was normal. Chromosomal microarray and Angelman, Prader-Willi and Fragile X testing were also normal.

Trio analysis of whole exome sequencing identified a *de novo* hemizygous nonsense variant in *TCEAL1*, c.259C>T, p.(Gln87*), which was absent from gnomAD and predicted pathogenic. Sanger

sequencing was not able to be performed but the variant was present in 25 of 25 reads by NGS in the proband and was not present in 35 and 44 reads in the parents.

Individual 4 was a 7-year-old male with hypotonia, hypermobility and moderate to severe developmental delay. He was born after an uncomplicated pregnancy as a second child to healthy parents. After an initial normal start, he presented with joint hypermobility, feeding difficulties and constipation within the first half a year. He was known to have a history of respiratory tract and ear infections. Examination at one year of age revealed hypotonia and motor developmental delay. Later examinations noted global developmental delay including speech and language delays. At 4 years old, he started walking, communicated mainly through gestures and some single words, and had developed some stereotypic movements and tics. He has an iris coloboma. He had a broad forehead but no other dysmorphic features and growth parameters were within normal range. Brain MRI showed delayed myelinization of the terminal zones of the lateral ventricles and slight diminishing of white matter in the parietooccipital zones.

Individual 5 (BAB13799; Family HOU5113) was a 17-year-old female with mild to moderate intellectual disability. She was born at 37 weeks via NSVD to a 30-year-old G2P0 mother and a 34-year-old father after an uncomplicated pregnancy. There were no reported prenatal exposures. Birth weight was 6 lbs. 10 oz. (SD -0.52), birth length was ~19 inches (SD -0.48), and birth head circumference is not recalled. Apgar scores were not recalled. The neonatal period was complicated by hyperbilirubinemia, requiring five days of inpatient phototherapy. Medical record review indicated a peak bilirubin level of 17 mg/dL.

There were concerns for developmental delay from infancy. The individual crawled at 13 months and walked at 22 months. The timing of her first word was not recalled, but at age two she had approximately 10 words. She had been in special education classes since age 2 years and received speech therapy for moderate/severe expressive and receptive language disorder, and occupational therapy to address hand and truncal weakness. She was diagnosed with autism at age 7 years, previously having been diagnosed with pervasive developmental disorder-not otherwise specified (PDD-NOS). She was noted to have vocal and motor tics in childhood.

The individual had a significant change in her neurobehavioral features at age 11 years. Approximately two weeks after a febrile illness characterized by myalgia and sore throat, the individual demonstrated acute changes in her behavior and mood as well as fatigue, facial and shoulder tics, intrusive thoughts and staring spells. An evaluation was conducted for infectious causes of her symptoms including streptococcal infection and Lyme disease. No evidence of recent infection was identified, however the neurobehavioral features subsided after approximately 6 weeks of treatment with clarithromycin, but later recurred. After their recurrence, the individual was treated with antibiotics, and later intravenous immune globulin and steroids and the immunomodulators mycophenolate or methotrexate. These episodes were responsive to steroids, but not prevented by immunomodulation.

In recent years, symptoms continued to wax and wane, at times dramatically, requiring approximately 3-4 oral steroid bursts per year. At 17 years old, she had an episode that did not resolve with oral steroids and has since been treated with low dose guanfacine and naltrexone as well as penicillin shots every 3-4 weeks. Additional infectious work up was performed and was unrevealing. Over the course of 4+ years of symptoms she was treated with multiple medications, including clozapine, desvenlafaxine, ziprasidone, sertraline, lithium, lorazepam, clonazepam and lurasidone with limited response if any. Non-prescription cannabidiol was helpful for relief of OCD symptoms (e.g., compulsive scribbling with pen on paper). Over time many of the individual's symptoms of fatigue and cognitive difficulties improved. However, there were persistent auditory hallucinations.

Clinical genetics evaluation included karyotype, fragile X testing, DNA microarray and trio-ES and GS. Additional evaluations included EEG, MRI of brain, and EKG. She had a history of microcytic anemia and is an alpha-thalassemia carrier. Eye exam showed slight astigmatism. Audiology screenings were normal. The individual was diagnosed with oral allergy syndrome triggered by most fresh fruits and vegetables. This individual had a history of hypertriglyceridemia.

Physical exam at age 14y and 16y revealed epicanthal folds, slight overbite, narrow hands with slender fingers, clinodactyly of the 4^{th} and 5^{th} toes, hallux valgus, mild shortening of the lateral toes, a

single café au lait spot on the back and an unsteady gait. Anthropometric measurements at age 16 years: OFC 57.5 cm, >98th centile (+2.32 SD); weight 62.4kg, 76th centile; and height 159.4 cm, 30th centile. The individual was of European and Chinese ancestry. She had a younger sister who was reportedly in good health. Family history was negative for close relatives with psychiatric disease or developmental disorders.

Review of the chromosomal microarray data identified a *de novo* heterozygous ~14 Kb deletion of ChrX:102,879,326-102,893,312 (hg19) with 77 bp of microhomology at the breakpoint site; *TCEAL1* was the only gene within the deleted region. This deletion was not present in the Database of Genomic Variants (http://dgv.tcag.ca/dgv/app/home) and was confirmed through further characterization of the breakpoints as described below.

Individual 6 was a 9-year-old female with severe intellectual disability and obesity. She was the fourth child to healthy Algerian non-consanguineous parents. She was born at 40 weeks of gestation. Birth weight was 3950 g (92th centile), birth length 51.5 cm (84th centile) and occipito-frontal circumference (OFC) 36 cm (87th centile). Neonatal period was marked by axial hypotonia. Psychomotor development was delayed: she sat at 12 months of age, walked unassisted at 19 months of age, and spoke her first words at 4 years of age. At 18 months of age, she had an accelerated weight and length gain with increase of the weight curve up to +5SD. She had surgeries at 7 years of age for bilateral genu valgum due to obesity. At 8 years of age fine motor skills were in the 20-month level, gross motor skills were in the 14-month level and expressive and receptive language were below a 12-month level. She had autistic spectrum disorder features (stereotypies, disturbances in social interactions) without a diagnosis of autism due to her severe intellectual disability. At last examination at 9 years of age, she spoke about ten words with some associations of words. Height was 136 cm (+2.5SD), weight 74 kg (>+5SD), OFC at 55.5 cm (+2.5SD) and body mass index at 40 (morbid obesity). She walked with small steps and open feet. Physical examination identified dysmorphic facial features including frontal bossing, bilateral epicanthus, hypertelorism, deep set eyes, horizontal eyebrows, fleshy earlobes and diastema. Ophthalmology examination including

electroretinogram showed astigmatism and strabismus. X-rays showed congenital brachymetatarsia involving the fourth metatarsal bone in the right foot. Head CT at 36 months of age and brain MRI at 4 years of age were normal. Metabolic workup including purines and pyrimidines, creatinine, urinary organic acids, mucopolysaccharides and oligosaccharides in the urine, as well as Fragile X testing, *GNAS* gene screening and a next-generation sequencing panel of cognitive deficiency genes were normal.

SNP array showed a 154 Kb *de novo* Xp22.2 deletion encompassing 3 genes: *TCEAL4, TCEAL3* and only one OMIM gene *TCEAL1* (*300237). No duplication encompassing *TCEAL1* was reported in DGV or gnomAD Structural Variant databases. Genome sequencing confirmed the Xq22.2 *de novo* deletion seq[GRCh37] Xq22.2(102,774,750_102,929,222)x1 *dn*. No additional pathogenic variants were identified, including all known genes involved in neurodevelopmental disorder.

Individual 7 was a 6-year-old female with global developmental delay, mild intellectual disability, profound speech delay and epilepsy. She was born at 38 weeks as a second child (third pregnancy, one spontaneous abortion) to unrelated and excluding refractory error in the mother, otherwise healthy parents via C-section due to the mother's myopia. Mother reported decreased fetal movements in comparison to the pregnancy with the previous child. Postnatal hypotonia was also noted. Both motor and intellectual development were delayed. The individual walked independently from the age of 18 months. Until the age of 13 months, she only crawled. Speech development was delayed: until the age of 3 years, she did not speak any words, but rather used non-verbal communication, a few onomatopoeic sounds, and proto-declarative behavior. Until the age of 5 years, her vocabulary consisted of less than 20 words. After the oxcarbazepine treatment initiation, her vocabulary range increased and she began to combine words into sentences. She reacted to others' emotions, liked to attract others' attentions, enjoyed physical contact (hugging), and reacted aggressively when ignored. She attended municipal kindergarten as well as rehabilitation classes. In the Stanford-Binet test, she was assessed as having mild intellectual disability. She also displayed self-stimulation, episodic temper tantrums and severely increased appetite. EEG revealed focal paroxysmal activity in the left frontocentrotemporal area with sharp waves and slow waves. After the

first atypical absence seizures at the age of 4 years, she was diagnosed with epilepsy and treated with oxcarbazepine with dose escalation to 150 mg – 0 – 450 mg. Good seizure control was achieved on this regimen. Head MRI at 6 years of age revealed a few small foci of increased signal in F2/FLAIR images, located in subcortical and periventricular areas of the white matter of both hemispheres and right dentate nucleus – with no evidence of contrast enhancement. Signs of upper respiratory tract inflammation was also noted: thickening of the maxillary sinus mucosa and enlargement of the pharyngeal tonsils. Slight facial dysmorphic features, including telecanthus, broad forehead, low-set ears and widely spaced teeth were also observed, and a single lentigo on the scalp. Other health problems included obesity, astigmatism and premature puberty (breast enlargement).

Methylation status of the Prader-Willi/Angelman critical region was normal (commercially available MS-MLPA Probemix ME028), previous *MAGEL2* sequencing for Schaaf-Yang syndrome revealed no pathogenic or potentially pathogenic variants. Singleton ES performed in the proband revealed a heterozygous ultra-rare (absent in gnomAD database) missense variant in *TMEM240* and frameshift variants in *ST6GALNAC6* and *TCEAL1*. With segregation studies, the *TMEM240* and *ST6GALNAC6* variants were found to be inherited from the proband's healthy father (the *TMEM240* variant was also observed in the proband's healthy brother, data not shown) and were classified as benign. However, we noted that, being extremely rare, the presence of these variants in the father confirmed parenthood. The frameshift variant in *TCEAL1*, c.169delC p.(Leu57Serfs*36), was absent in both parents and the proband's brother, thus it was considered as a *de novo* event.

Individual 8 was a 12-year-old male of European ancestry who came to medical attention for features of spastic paraparesis. He was born to a non-consanguineous couple who were 35 (father) and 36 (mother) years of age at term, and there were no immediate concerns during the neonatal period. Birth weight was 3480 g (44.1%, -0.15 SD). He met his early developmental milestones during the first year of life. At the age of 13 months, he began walking independently, and his gait was described as idiopathic toe-walking of childhood. He would walk only short distances, and began to crawl increasingly. Sometime after he began

walking, lower extremity spasticity was noted, with additional difficulties in fine motor skills involving the upper extremities. By the age of 8 years, he had mild dysarthria when speaking English (his second language), no dysphagia, full extra-ocular movements, normal visual acuity and auditory functioning, and presumed normal intellect based on his performance in the third grade. He did not demonstrate any bowel incontinence, but did develop urinary incontinence. His motor exam was notable for 4+/5 weakness involving the shoulders and lower extremity flexors with evidence for spasticity of the lower extremities more severely than the upper extremities. He was hyperreflexic with an upgoing Babinski, and a normal sensory exam. Rapid alternating movements were slow with his upper extremities, but there was no dysmetria. His gait was very spastic. He was also noted to have mild scoliosis. Clinical investigation included a normal brain MRI, normal CT scan of the head, and normal EMG.

His family history was notable for two parents in good health, and two older sisters, ages 12y and 10y, who were also reported to be in good health. Neurodegenerative conditions were reported on both maternal and paternal sides of the family: progressive supranuclear palsy (PSP) in the maternal grandmother (deceased beyond 60 years of age from the same), Parkinson disease (PD) in the paternal grandfather (deceased at 42 years of age from the same), and one maternal cousin who is presently living with multiple sclerosis (MS).

Review of exome data identified a hemizygous missense variant in *TCEAL1*, c.346G>A p.(Asp116Asn) (Figure 3), which was present in gnomAD in one female European individual with a population specific allele frequency of 0.000012 and an overall allele frequency of 0.00000545. The variant was predicted to be deleterious by SIFT but benign by PolyPhen-2 and had a CADD phred score of 24.1 and a REVEL score of 0.139 (Table 1, Table S2). The variant was found to be inherited from the unaffected mother. Further analysis of the ES data by ExomeDepth did not reveal any pathogenic copy number variants.

SUPPLEMENTAL TABLES

Table S2 Detailed Molecular Information

	Individual 1	**Individual 2**	**Individual 3**	**Individual 4**	**Individual 5**	**Individual 6**	**Individual 7**	**Individual 8**
Sex	M	M	M	M	F	F	F	M
Ancestry	White (U.S.)	White (European)	White (European)	White (European)	White + Chinese	Algerian	White (European)	White (European)
Consanguinity	No	No	No	No	No	No	No	No
Age at most recent clinical evaluation	5 y/o	10 y/o	17 y/o	7 y/o	17 y/o	8 y/o	6 y/o	12 y/o
Variant type	SNV Nonsense	Indel Frameshift	SNV Nonsense	SNV (Missense)	CNV (single gene) Deletion	Xq22 deletion	SNV Frameshift	SNV Missense
Variant	NM_001006639: c.447G>A p.(Trp149*)	NM_001006639: c.299_302del p.(Gly100Alafs*22)	NM_001006639: c.259C>T p.(Gln87*)	NM_001006639: c.269G>A p.(Cys90Tyr)	Single-gene deletion of *TCEAL1*	Deletion of *TCEAL1, TCEAL3, TCEAL4*	NM_001006639: c.169delC p.(Leu57Serfs*36)	NM_001006639: c.346G>A p.(Asp116Asn)
Zygosity	Hemizygous 39 of 39 reads Sanger confirmed	Hemizygous 12 of 12 reads	Hemizygous 25 of 25 reads	Hemizygous Sanger confirmed	Heterozygous	Heterozygous	Heterozygous 42 out of 79 reads	Hemizygous Sanger confirmed
Coordinates (hg19)	ChrX: 102,885,291	ChrX: 102,885,138	Chr X: 102,885,103†	ChrX: 102,885,113	ChrX: 102,879,326- 102,893,312	ChrX: 102,774,750- 102,929,222	ChrX: 102,885,012	ChrX: 102,885,190
Inheritance	*De novo* Sanger confirmed	*De novo* by NGS	*De novo* by NGS	*De novo* Sanger confirmed	*De novo* HD-aCGH confirmed	*De novo* SNP-array, WGS	*De novo* NGS-based amplicon deep sequencing	Maternally inherited Sanger confirmed
Population frequency	NP (gnomAD)	NP (gnomAD)	NP (gnomAD)	NP (gnomAD)	NP (DGV, DECIPHER, gnomAD SV)	NP (DGV, gnomAD)	NP (gnomAD)	1/81,897 European alleles (gnomAD)
In silico missense predictions	Truncation	Truncation	Truncation	SIFT: Tolerated (0.09) Poly-Phen-2: Probably damaging (0.997)	Full gene deletion	Full gene deletion	Truncation	SIFT: Deleterious (0.1) PolyPhen-2: benign (0.429)

CADD and REVEL scores	37; NA	NA	36; NA	23.8; 0.146	NA	NA	NA	24.1; 0.139
Genetic testing performed	Trio ES	Trio ES	Trio ES	Trio ES	Trio-CMA (BG) Trio HD-aCGH	Trio GS Trio SNP array	Singleton ES	Singleton ES
Other genetic/ environmental factors contributing to phenotype	None	None	None	None	Carrier for alpha thalassemia trait	None	None	None
Other variants identified but interpreted as probably non-pathogenic	None known/reported	None known/reported	None known/reported	Two regions of homozygosity (7.77 Mb and 19.38 Mb) with no known clinical significance	Het SNV in *HAAO* (NM_012205.3: c.685T>C p.(W229R)) - AR disease (MIM# 617660) mainly skeletal/non-neurological	None known/reported	*TMEM240* NM_001114748.2: c.481A>C p.(Lys161Gln) inherited from a healthy father	None known/reported

BG: Baylor Genetics, CADD: Combined Annotation Dependent Depletion, CMA: chromosomal microarray analysis, HD-aCGH: high definition array comparative genomic hybridization, CNV: copy number variant, DGV: database of genomic variants, ES: exome sequencing, F: female, GS: genome sequencing, M: male, NA: not applicable, NP: not present; REVEL: rare exome variant ensemble learner, SNV: single nucleotide variant, WGS: whole genome sequencing. †converted from hg38 coordinates (chrX:103,630,175).

NM_001006639.2 is used for variant nomenclature.

Table S3 Amino Acid identity of TCEAL proteins with TCEAL1

The TCEAL protein family includes 9 proteins (TCEAL1, TCEAL2, TCEAL3, TCEAL4, TCEAL5, TCEAL6, TCEAL7, TCEAL8, and TCEAL9) located in Xq22.1-Xq22.2. The percent amino acid identity of each protein with TCEAL1 is shown.

Gene	**Genomic Region**	**% amino acid identity with TCEAL1**
TCEAL2	Xq22.1	42.11%
TCEAL3	Xq22.2	No significant similarity
TCEAL4	Xq22.2	31.25%
TCEAL5	Xq22.1	No significant similarity
TCEAL6	Xq22.1	No significant similarity
TCEAL7	Xq22.1	35.87%
TCEAL8	Xq22.1	52.83%
TCEAL9	Xq22.2	53.33%

SUPPLEMENTAL METHODS

GENOMIC METHODS

Genomic methods for each subject are summarized in Supplemental Table 2. **Individuals 1-4 and 7-8** underwent exome sequencing (ES) or genome sequencing (GS). ES in **individual 1** and his parents was completed through Psomagen (Rockville, MD), with trio analysis completed using SNP & Variation Suite (Golden Helix, Bozeman, MT) including gnomAD tracts for general population frequency and dbNSFP for *in silico* predictions as described.[1] Query of the MatchMaker exchange identified seven additional individuals hosted by MyGene2[2], GeneMatcher[3,4] and DECIPHER.[5,6] For **individual 2**, ES and trio analysis were completed through the Deciphering Developmental Disorders project, and genome sequencing (GS) and trio analysis were completed through the 100,000 genomes project. For **individual 3**, clinical ES and trio analysis was performed in the Genetics department of Angers University Hospital using a Twist Biosicence (San Francisco, CA) in-solution enrichment methodology (Twist Human Core Exome EF Multiplex + Twist Human RefSeq Panel), followed by 101 bases of paired-end reads via a massively parallel sequencing approach on Illumina NextSeq550. Bcl were converted as FASTQ using GenerateFASTQ (v.1.0.0). Quality control was done using Fastqc (v.0.11.9). Sequence reads were mapped to the human genome build (hg38 / GRCh38) and analyzed with a dedicated in-house pipeline integrating various modules for coverage analysis, variant calling, annotation. Variant prioritization was performed with VarAFT 2.16.[7] For **individual 4**, clinical trio ES was performed using SureSelect V5 (Agilent Technologies) capture reagent, and sequenced on an Illumina HiSeq 2500. Coding and splice site exome variants were prioritized based on predicted *de novo* or autosomal recessive disease transmission and following the diagnostic standards of the University Medical Center Utrecht. **individual 7** underwent singleton clinical ES performed using Twist Human Core Exome EF Multiplex + Twist Human RefSeq Panel (Twist Bioscience) and paired-end (2x100 bp) sequenced on Illumina NovaSeq 6000. Family study was performed by next generation sequencing-based deep amplicon sequencing. **individual 8** underwent singleton ES with CNV analysis using ExomeDepth[8].

For all SNVs, i*n silico* analysis included SIFT and Polyphen-2 predictions of the effect of missense variants as well as CADD[9] and REVEL[10] scores. Variant segregation within families by PCR and subsequent Sanger sequencing was performed for two of the four ES-identified *TCEAL1* SNVs. Primer pairs were used for the segregation PCRs as follows:

Individual	**Forward Primer (5'-3')**	**Reverse Primer (5'-3')**
Individual 1	AGCAGCCTCCTTGTGGAGTA	ATCTTTCATGCAAATGTGTAGGGC
Individual 4	TGAGGAGCTCTTGCCTGA	GTATTGCCATTGCCACCT
Individual 8	CCTCGGAGGAGGAGTTCTTT	CAGAAAGTCCAGGTGGCAAT

PCR and Sanger sequencing was not possible for individual 2 (independent exome sequencing and genome sequencing trio analyses identified the *de novo* variant), individual 3 (variant was present in 25/25 NGS reads and not present in either parent), or individual 7 (variant was present in 3716 reads out of 8061 NGS-base amplicon deep sequencing reads and not present in either parent or healthy brother, with depth of coverage at 13,897x, 7514x, and 13084x, respectively).

Individual 5 was identified through a screen for single-gene CNVs in the clinical microarray database of the Baylor Genetics (BG) diagnostic laboratory (consisting of > 75,000 clinical genomes). The deletion in individual 5 was initially detected by clinical-grade Chromosomal Microarray Analysis (CMA; version 11.2) at BG. This CMA version has a total of 400k probes spread genome-wide at alternating average densities of '> 4 probes per exon' in >4,200 clinically-relevant genes and '1 probe per 30 Kb' genome-wide in non-coding regions.[11] Due to the lower probe resolution of CMA array in non-coding regions, a higher resolution custom array was utilized for the precise mapping of deletion breakpoints; this high-density (HD) array was designed on the Agilent eArray website (https://earray.chem.agilent.com/suredesign/) and interrogates copy-number variation (CNV) in Xq22 via a total of 40,208 60-mer probes spanning chrX:97,915,511–113,400,000 (NCBI build 36) at an average distribution of one probe per 386 bp (format 4 × 44 K).[12] It was utilized to assay the genomes of this subject (individual 5, BAB13799), and her parents

(BAB13800 and BAB13801). The experimental procedures were performed according to the manufacturer's protocol (Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis, Version 7.2) with some modifications.[13] Sex-matched controls (GM15510 as female control and GM10851 as male control, both from the Coriell Institute) were used with the trio. The scanned array images were processed by the Agilent Feature Extraction software (version10) and the extracted files were analyzed by the Genomic Workbench (version 7.0.4.0). Maximum proximal and distal boundaries of the heterozygous X-linked deletion were defined as follows, proximally as the last probe with a log2 ratio of 1 prior to the deleted probes (avg. log2 ratio of −1) and distally as the first probe with a log2 ratio of 1 after the deleted probes.

Precise mapping of the deletion breakpoints to the nucleotide level was accomplished by PCR and Sanger sequencing using inward facing primers that completely span the deletion in individual 5 (i.e. primers that bind outside of the maximum boundary coordinates deduced from HD-aCGH results). Parental samples were studied alongside the proband's to confirm specificity of the PCR product in the proband. The Jct-PCR primer-pair sequences are as follows, F (5'->3'): GAGGCCTTGCATTGTCTTTT and R (5'->3'): AAAGCTGCTCCCAGGAAAGT. The obtained sequence (i.e. the deletion breakpoint-junction sequence or Jct-sequence) was aligned to the haploid reference genome (GRCh37/hg19) and breakpoint mapping coordinates were evaluated for overlapping repeat sequences using four different repeat datasets, 1) the highly similar intrachromosomal repeats (HSIR) dataset,[14] 2) the Segmental Dup dataset,[15,16] 3) the Repbase (Repeat Masker) dataset,[17] and 4) the SelfChain dataset.[18,19]

Individual 6 underwent a parent-proband clinical trio-SNP array (Illumina OmniExpress 700K) as well as trio-genome sequencing for fine-mapping of breakpoints. X chromosome inactivation (XCI) studies were performed clinically by analysis of methylation status at the following loci: *PCSK1N* (Xp11.23), *AR* (Xq11.2-2q12), *ZDHHC15* (Xq13.3), and *SLITRK4* (Xq27.3).

INFORMED CONSENT

Written informed consent for research studies and/or reporting of clinical features was obtained for all Individuals, including photo publication if applicable; additional clinical details of these individuals are provided herein. Human studies were approved by the Institutional Review Boards of Children's Wisconsin (protocol 124172) (individual 1), UK Research Ethics Committee (REC) (protocol 10/H0305/83 granted by the Cambridge South REC and GEN/284/12 granted by the Republic of Ireland REC) and Health Research Authority REC (individual 2), Baylor College of Medicine (protocol H-29697) (individual 5), and McGill University Health Centre (protocol 31-01-008703) (individual 8); the remaining variants were identified through clinical testing with informed consent to publish. Genomic DNA for all experiments was isolated from blood according to standard procedures. All available neuroimages were reviewed by an expert board-certified neuro-radiologist (JVH).

SUPPLEMENTAL ACKNOWLEDGEMENTS

This study makes use of data generated by the DECIPHER community and the DDD study.

The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network.

This study makes use of data generated by the DECIPHER community. A full list of centres who contributed to the generation of the data is available from https://deciphergenomics.org/about/stats and via email from contact@deciphergenomics.org. Funding for the DECIPHER project was provided by Wellcome. Those who carried out the original analysis and collection of the Data bear no responsibility for the further analysis or interpretation of the Data.

SUPPLEMENTAL REFERENCES

1. Reis, L.M., Sorokina, E.A., Thompson, S., Muheisen, S., Velinov, M., Zamora, C., Aylsworth, A.S., and Semina, E.V. (2019). *De Novo* Missense Variants in *WDR37* Cause a Severe Multisystemic Syndrome. Am J Hum Genet *105*, 425-433.
2. MyGene2. NHGRI/NHLBI University of Washington-Center for Mendelian Genomics, Seattle, WA (https://mygene2.org/MyGene2/). http://www.mygene2.org.
3. Azzariti, D.R., and Hamosh, A. (2020). Genomic Data Sharing for Novel Mendelian Disease Gene Discovery: The Matchmaker Exchange. Annu Rev Genomics Hum Genet *21*, 305-326.
4. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat *36*, 928-930.
5. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet *84*, 524-533.
6. Swaminathan, G.J., Bragin, E., Chatzimichali, E.A., Corpas, M., Bevan, A.P., Wright, C.F., Carter, N.P., Hurles, M.E., and Firth, H.V. (2012). DECIPHER: web-based, community resource for clinical interpretation of rare variants in developmental disorders. Hum Mol Genet *21*, R37-44.
7. Desvignes, J.P., Bartoli, M., Delague, V., Krahn, M., Miltgen, M., Beroud, C., and Salgado, D. (2018). VarAFT: a variant annotation and filtration system for human next generation sequencing data. Nucleic Acids Res *46*, W545-W553.
8. Plagnol, V., Curtis, J., Epstein, M., Mok, K.Y., Stebbings, E., Grigoriadou, S., Wood, N.W., Hambleton, S., Burns, S.O., Thrasher, A.J., et al. (2012). A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. Bioinformatics *28*, 2747-2754.
9. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., and Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res *47*, D886-D894.
10. Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., et al. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet *99*, 877-885.
11. Gambin, T., Yuan, B., Bi, W., Liu, P., Rosenfeld, J.A., Coban-Akdemir, Z., Pursley, A.N., Nagamani, S.C.S., Marom, R., Golla, S., et al. (2017). Identification of novel candidate disease genes from *de novo* exonic copy number variants. Genome Med *9*, 83.

12. Carvalho, C.M., Bartnik, M., Pehlivan, D., Fang, P., Shen, J., and Lupski, J.R. (2012). Evidence for disease penetrance relating to CNV size: Pelizaeus-Merzbacher disease and manifesting carriers with a familial 11 Mb duplication at Xq22. Clin Genet *81*, 532-541.
13. Carvalho, C.M., Zhang, F., Liu, P., Patel, A., Sahoo, T., Bacino, C.A., Shaw, C., Peacock, S., Pursley, A., Tavyev, Y.J., et al. (2009). Complex rearrangements in patients with duplications of *MECP2* can occur by fork stalling and template switching. Hum Mol Genet *18*, 2188-2203.
14. Hijazi, H., Coelho, F.S., Gonzaga-Jauregui, C., Bernardini, L., Mar, S.S., Manning, M.A., Hanson-Kahn, A., Naidu, S., Srivastava, S., Lee, J.A., et al. (2020). Xq22 deletions and correlation with distinct neurological disease traits in females: Further evidence for a contiguous gene syndrome. Hum Mutat *41*, 150-168.
15. Bailey, J.A., Gu, Z., Clark, R.A., Reinert, K., Samonte, R.V., Schwartz, S., Adams, M.D., Myers, E.W., Li, P.W., and Eichler, E.E. (2002). Recent segmental duplications in the human genome. Science *297*, 1003-1007.
16. Bailey, J.A., Yavor, A.M., Viggiano, L., Misceo, D., Horvath, J.E., Archidiacono, N., Schwartz, S., Rocchi, M., and Eichler, E.E. (2002). Human-specific duplication and mosaic transcripts: the recent paralogous structure of chromosome 22. Am J Hum Genet *70*, 83-100.
17. Jurka, J. (2000). Repbase update: a database and an electronic journal of repetitive elements. Trends Genet *16*, 418-420.
18. Chiaromonte, F., Yap, V.B., and Miller, W. (2002). Scoring pairwise genomic sequence alignments. Pac Symp Biocomput, 115-126.
19. Schwartz, S., Kent, W.J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R.C., Haussler, D., and Miller, W. (2003). Human-mouse alignments with BLASTZ. Genome Res *13*, 103-107.