Mutation in an alternative transcript of *CDKL5* in a boy with early-onset seizures

Dale L. Bodian,[1] John M. Schreiber,[2] Thierry Vilboux,[1] Alina Khromykh,[1] and Natalie S. Hauser[1]

[1]Inova Translational Medicine Institute, Inova Health System, Falls Church, Virginia 22042, USA; [2]Pediatric Specialists of Virginia, Falls Church, Virginia 22042, USA

Abstract Infantile-onset epilepsies are a set of severe, heterogeneous disorders for which clinical genetic testing yields causative mutations in ∼20%–50% of affected individuals. We report the case of a boy presenting with intractable seizures at 2 wk of age, for whom gene panel testing was unrevealing. Research-based whole-genome sequencing of the proband and four unaffected family members identified a de novo mutation, NM_001323289.1: c.2828_2829delGA in *CDKL5*, a gene associated with X-linked early infantile epileptic encephalopathy 2. *CDKL5* has multiple alternative transcripts, and the mutation lies in an exon in the brain-expressed forms. The mutation was undetected by gene panel sequencing because of its intronic location in the *CDKL5* transcript typically used to define the exons of this gene for clinical exon-based tests (NM_003159). This is the first report of a patient with a mutation in an alternative transcript of *CDKL5*. This finding suggests that incorporating alternative transcripts into the design and variant interpretation of exon-based tests, including gene panel and exome sequencing, could improve the diagnostic yield.

[Supplemental material is available for this article.]

INTRODUCTION

Epilepsies are one of the more common groups of neurological syndromes. These heterogeneous seizure disorders have diverse etiologies and a broad range of age of onset and severity, with the most severe forms beginning in infancy. Among children of <2 yr of age, the incidence of epilepsies is estimated at 70 per 100,000 (Eltze et al. 2013). Many epilepsies have a genetic basis, and more than 500 genes have been associated with these disorders (Weber et al. 2017).

Clinical gene panel sequencing with copy-number variant (CNV) analysis can reveal candidate disease-causing mutations for <20% to >50% of severe epilepsy patients, with the highest yields obtained for sporadic, early-onset cases (Trump et al. 2016; Sands and Choi 2017). Negative results from gene panel testing have been attributed to nongenetic disease etiology, mutation of genes not included in the panels, technical limitations in detecting CNVs or somatic mutations, and difficulty in interpreting variant pathogenicity (Sands and Choi 2017). Genes such as the seizure-related gene *CDKL5* undergo alternative splicing, but not all exons are necessarily represented on the gene panels, providing another possible explanation for negative results. Authors of previous studies characterizing the alternative transcripts of *CDKL5* recommended screening patients for variants in these additional exons,

but to date no patient with a mutation affecting an alternative isoform of CDKL5 has been reported (Fichou et al. 2011; Rademacher et al. 2011; Williamson et al. 2012; Hector et al. 2016); hence, the clinical relevancy of mutations in these exons has not yet been established.

We present the case of a severely affected boy with intractable early-onset seizures for whom clinical genetic testing was unrevealing. Whole-genome sequencing (WGS) of the family revealed a de novo mutation in *CDKL5* that is exonic in brain-expressed transcripts but not in the testis-expressed transcript typically chosen as the reference transcript for this gene in epilepsy gene panel tests and in clinical exome sequencing. This result supports the proposal that alternative exons, in *CDKL5* as well as other genes, should be screened in clinical sequencing and demonstrates that the choice of reference transcripts used in assay design and variant interpretation can lead to overlooked causative mutations even within genes represented in the test.

RESULTS

Clinical Presentation and Family History

The patient was born full term without complications. Tonic seizures began at 2 wk of life, characterized by symmetric tonic arm flexion or extension, lasting ~30–60 sec, averaging 10 per day. At age 7 mo, he developed infantile spasms. Antiseizure therapies were initiated and trialed in this order: levetiracetam, topiramate, clonazepam, oxcarbazepine, high-dose prednisolone, the ketogenic diet, clobazam, cannabidiol, valproic acid, ACTH, and rufinamide. After starting oxcarbazepine at 7 mo, the electroencephalogram (EEG) background deteriorated significantly (see below), coincident with the onset of infantile spasms. He now receives levetiracetam, valproic acid, and clobazam and is restarting the ketogenic diet. All other antiseizure medications were discontinued because of lack of efficacy. Currently, seizures are generalized tonic and focal seizures with alteration in awareness. Development has never progressed beyond a 1-mo-old level; the patient occasionally smiles and regards and exhibits profound hypotonia, intellectual disability, and cortical vision impairment. Head circumference, length, and weight have remained normal. His medical and surgical history is also notable for G-tube dependence, chronic lung disease, and Nissen fundoplication.

His EEGs (Supplemental Fig. S1a–c) typically revealed multifocal epileptiform activity, at first predominantly in the posterior head regions. Background activity was initially reported to be within the broad range of normal limits for his age, but became disorganized, with high-amplitude delta activity, and multifocal epileptiform discharges, consistent with hypsarrhythmia at 7 mo old. Magnetic resonance imaging (MRI) brain at 2 mo old was normal for his age. Repeat MRI brain at 25 mo demonstrated mild diffuse volume loss. Testing for inborn errors of metabolism including lactic acid, pyruvic acid, ammonia, plasma amino acids, urine organic acids, serum acylcarnitine profile, free and total carnitine, uric acid, biotinidase, urine *S*-sulfocysteine, cerebrospinal fluid routine studies, and neurochemistry were all normal. A muscle biopsy showed normal electron microscopy with no abnormalities. Screenings for congenital disorders of glycosylation, neuronal ceroid lipofuscinosis, and peroxisomal disorders were normal. A whole-genome CGH (GenomeDx v5, GeneDx) and the GeneDx Infantile Epilepsy Panel (sequencing and deletion/duplication analysis of 38 genes) were reported as normal.

The family history is fairly remarkable. The older sister, now neurotypical, had a history of hypotonia, motor delay, and one febrile seizure. An older brother is neurotypical. His mother has a history of fibromyalgia, anxiety, and atherosclerosis. His father had febrile seizures that did not persist after childhood. The paternal grandfather also had a history of seizures associated with an episode of a high fever. One maternal aunt had multiple miscarriages, two associated with brain anomalies. There is no known consanguinity. Although the family

history includes three members with histories of febrile seizures, none of the seizure episodes lasted past young childhood, and all are healthy.

Genomic Analyses

To identify candidate explanatory variants in a wider range of genomic regions than were assessed by the clinical epilepsy panel, the proband and his family were enrolled in Inova Translational Medicine Institute's (ITMI's) IRB-approved research study entitled "The Incidence and Burden of Congenital Anomalies, Genetic Disorders and Genetic Suspicion in Neonatal and Pediatric Patients." WGS was performed at 40× coverage (Supplemental Table S1) on peripheral blood samples from the proband, both parents, an unaffected brother, and an unaffected sister (Fig. 1). The 6,775,467 small variants and 3844 CNV in the family were computationally prioritized by predicted effect on the encoded protein, inheritance pattern, allele frequency, call quality, and reliability of the genomic region for sequencing. Variants within or overlapping genes previously associated with seizure disorders were flagged but the results were not limited to known epilepsy genes.

The automated filtering yielded two candidate variants, a de novo mutation NM_001323289.1:c.2828_2829delGA in the gene *CDKL5* and a maternally inherited variant NM_024490.3:c.509G>A in the gene *ATP10A* (Table 1). *ATP10A*, also known as *ATP10C,* is a maternally expressed gene located within the most common deletion associated with Angelman syndrome, 15q11-q13 (Meguro et al. 2001). However, the role of *ATP10A* in Angelman syndrome is not yet clear and no missense mutations in this gene have been reported in patients with seizures. The *ATP10A* variant in this family has been observed in reportedly healthy individuals with a maximum allele frequency of 0.0029, in the 1000 Genomes Project's Hispanic superpopulation (Auton et al. 2015). In Angelman syndrome, it is unusual for seizures to start in the neonatal period; median seizure onset is 2 yr of age, and the EEG often shows high-voltage slow waves or persistent rhythmic theta initially which evolves to triphasic or "notched" delta. Patients usually ambulate, walking with an ataxic gait, and do not typically have cortical visual impairment. They have a distinct facial
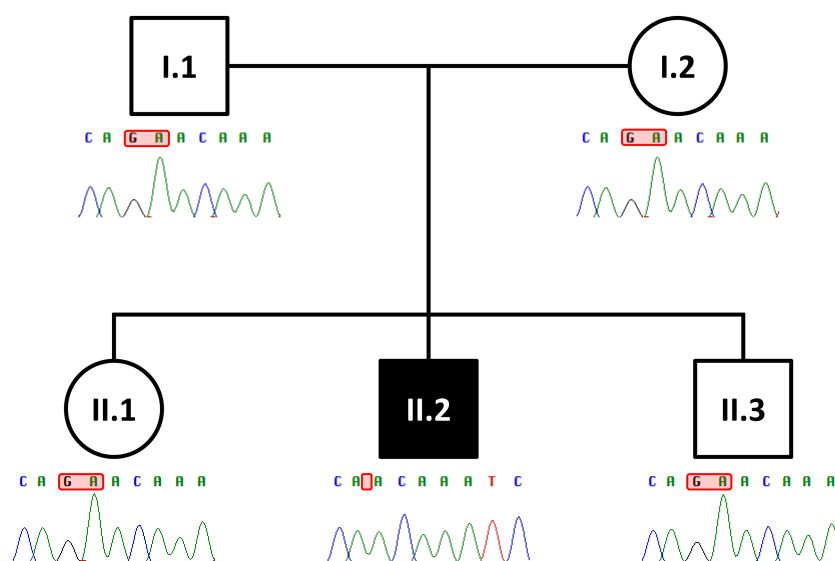

Figure 1. Family pedigree and Sanger sequencing of the region of the *CDKL5* variant. The 2-bp deletion NM_001323289.1:c.2828_2829delGA on Chromosome X was detected in the proband (black square) in the haploid state but not in the four unaffected family members.

Table 1. Candidate variants

Genomic (hg19)	cDNA	Protein	Gene	Maximum allele frequency[a]	ClinVar	rsid	Proband genotype	# Reads ref, alt[b]: proband mother father sister brother	Parent of origin
Chr X:18646820 CAG>C	NM_001323289.1: c.2828_2829delGA	p.R943fs	CDKL5	0	SCV000599451	–	Haploid	0,6 29,0 15,0 16,0 14,0	De novo
Chr 15:26026311 C>T	NM_024490.3: c.509G>A	p.R170H	ATP10A	0.0029	–	rs146522541	Heterozygous	11,18 17,13 26,0 31,0 43,0	Maternal

ref, reference allele; alt, alternate allele.

[a]Maximum allele frequency was computed from the gnomAD (Lek et al. 2016) and 1000 Genomes Project (Auton et al. 2015) subpopulation whole-genome data sets.

[b]Number of mapped reads in the WGS data supporting each allele for each individual.

appearance including a wide mouth, prognathism, and widely spaced teeth. Their behavior includes spontaneous laughter and frequent smiling (Galván-Manso et al. 2005). Our patient's presentation at such a young age of 2 wk of life, his EEG pattern, and the lack of the typical dysmorphic features and behavior seen in Angelman syndrome suggest that this variant is unlikely to be the primary cause of his condition, although we cannot rule it out as a contributing factor.

The second candidate variant lies in *CDKL5*, a gene associated with infantile epileptic encephalopathy 2 (EIEE2; OMIM:300672). This de novo 2-bp deletion lies in exon 17 of a transcript (NM_001323289) encoding the predominant isoform of CDKL5 expressed in human brain (Williamson et al. 2012; Hector et al. 2016). (In the nomenclature proposed by Hector et al. [2016], this mutation is located in the versions of exon 19 that contribute to isoforms hCDKL5_1-4.) Features associated with EIEE2, namely seizure onset in the first 3 mo of life, with severe developmental delay from birth and no overt dysmorphisms (Fehr et al. 2013), are observed in the proband. There may also be, as seen in our proband, associated gastrointestinal and respiratory problems, in addition to scoliosis that develops later (Mangatt et al. 2016). Because of the congruence of the patient's symptoms to those of previously reported *CDKL5*-associated EIEE2 patients, the de novo occurrence of the mutation, its absence from databases of whole-genome sequences, and its likely detrimental impact on the encoded protein, we consider this variant to be causative for the proband's condition. The presence of the mutation in the proband and its absence in the four unaffected family members was confirmed by research-based Sanger sequencing (Fig. 1). The mutation in the proband was further confirmed by clinical-grade targeted genetic testing via Sanger sequencing of the mutation of interest (GeneDx) from a newly collected buccal sample.

This *CDKL5* mutation identified by WGS was not found by the clinical epilepsy panel, even though the gene is screened in this test. To reconcile this difference, we compared the reference transcript(s) used to define the *CDKL5* exons for each analysis. According to the clinical report, and confirmed by the company, the panel screens the exons of the NM_003159 transcript, in which the patient's variant is intronic (NM_003159.2:c.2713+ 115delGA) (Fig. 2). WGS identified the mutation as exonic by examining alternative
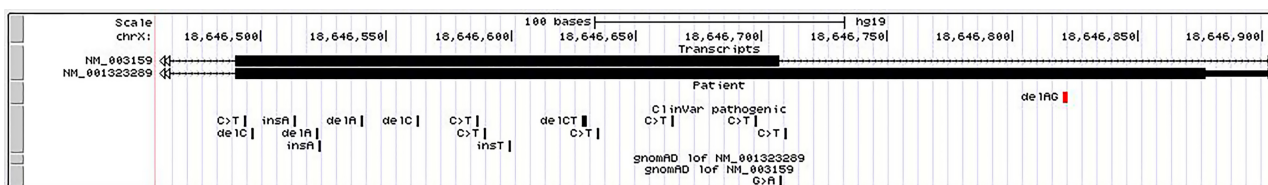
Figure 2. Genomic region of the *CDKL5* variant. Predicted coding sequence in this region from transcripts NM_003159 and NM_00132389 are represented by thick black bars, UTRs by thin bars, and intronic sequence by lines with arrowheads indicating the direction of transcription (Kent et al. 2002). The patient's mutation, labeled delAG, is indicated in red. ClinVar (Landrum et al. 2016) variants annotated as pathogenic or likely pathogenic are displayed. "gnomAD lof" tracks show variants from gnomAD (Lek et al. 2016) predicted to be stop-gain, frameshift, or splicing variants for NM_003159 or NM_00132389.

transcripts including NM_001323289. Interestingly, we contacted the six companies from which our hospital most often orders genetic tests (Ambry Genetics, Baylor, Fulgent, GeneDx, Invitae, and PreventionGenetics), and all use NM_003159 as the only reference transcript for *CDKL5* in both their gene panel and exome tests.

The patient's mutation lies in the 3′ region of NM_001323289 exon 17, whereas the gene panel transcript NM_003159 shares the 5′ region of this exon, but lacks the 3′ region (Fig. 2; Williamson et al. 2012; Hector et al. 2016). Pathogenic and likely pathogenic mutations in the shared 5′ region of this exon have been reported (Fig. 2). No pathogenic or benign mutations in the 3′ region beyond the flanking sequence of the NM_003159 exon are reported in ClinVar, consistent with the use of NM_003159 as the reference transcript for clinical sequencing. No predicted loss-of-function variants (stop-gain, frameshift, or splicing) in the NM_00132389 exon were found in the gnomAD database of genome sequences from healthy individuals (Lek et al. 2016), whereas one predicted loss of function variant, c.2713+1G>A, was found in one individual for NM_003159 (Fig. 2). This is consistent with NM_001323289 rather than NM_003159 being the disease-relevant transcript.

DISCUSSION

Gene panel sequencing has proved useful in aiding the diagnosis of patients with epilepsy of genetic etiology. However, even among the most likely molecularly diagnosed group of sporadic, early-onset cases, these tests are unrevealing for a subset of patients. As described previously, there are multiple reasons for negative results, ranging from nongenetic etiology to technical issues (Sands and Choi 2017). The mutation in the proband presented here is, to our knowledge, the first reported mutation in an exon of *CDKL5* that is not typically represented on clinical gene panels or exome tests, and it demonstrates that the choice of reference transcripts can also impact the diagnostic yield.

Although there are no functional data available confirming the pathogenicity of this mutation in the proband, the diagnosis of *CDKL5*-associated EIEE2 is strongly supported by the patient's clinical features, as well as the apparently de novo origin. The family history includes three relatives with febrile seizures; however, the proband's condition is not likely to result from a shared inherited variant that also causes febrile seizures in the relatives, as there generally is not overlap between the specific mutations that cause febrile seizures (in *SCN1A*, *SCN2A*, *SCN1B*, *GABRA1*, etc.) and those causing more severe epileptic encephalopathies. The clinical presentation and timing of resolution of the family members' symptoms strongly favors benign febrile seizures of childhood.

CDKL5-associated EIEE2 is an X-linked disorder for which boys are usually more severely affected than girls, and comprise <10% of reported cases (Bahi-Buisson and Bienvenu 2012).

The reasons for the skewed gender ratio are unknown. Reports of affected cases include boys with somatic mutations (Masliah-Plachon et al. 2010; Mei et al. 2014) or with a 47,XXY karyotype (Sartori et al. 2009), raising the possibility that most haploid mutations in males are lethal. However, as with our patient, other affected boys have a normal male karyotype and no evidence of mosaicism, with the reported mutations occurring in multiple exons of the brain-expressed transcripts (RettBASE: RettSyndrome.org Variation Database mecp2.chw.edu.au/). This report contributes a novel mutation and clinical data that can be used in the analysis of genotype–phenotype relationships of *CDKL5*-associated disorders.

CDKL5 has multiple transcripts, but the clinical significance of mutations in the alternative exons is unknown. Mutations in exons 19–21 of the commonly used reference transcript NM_003159, which encode the carboxy-terminal region of the CDKL5 isoform expressed predominantly in testis (Williamson et al. 2012; Hector et al. 2016), are unlikely to be pathogenic for EIEE2 (Diebold et al. 2014). In contrast, the mutation in the proband lies in an exon encoding the alternate carboxy-terminal region of the predominant brain isoforms, which have an expression pattern consistent with the disorder (Hector et al. 2016). The identification of this mutation supports the hypothesis that mutations in alternative exons of this gene can be clinically relevant. The lack of previously observed mutations in the 3′ region of this exon is likely due in part to the limited number of patients screened in this region, because it is not represented in commonly used gene panel and exome-sequencing tests.

The implication that mutations in an alternative exon of *CDKL5* can be clinically important should be considered when selecting appropriate genetic tests for patients suspected of having a *CDKL5*-related disorder. Patients who receive negative results from tests limited to the *CDKL5* NM_003159 transcript can consider a test that assays alternative exons in this gene. Our results suggest that inclusion of alternate reference transcripts in the design and analysis of exon-based assays aimed at identifying coding variants in *CDKL5*, including gene panel, exome, and genome sequencing, could lead to molecular diagnoses for more patients.

METHODS

Samples and WGS

DNA was extracted from peripheral blood and sent to Macrogen Clinical Laboratory for library preparation and WGS. Libraries were prepared using the Illumina TruSeq DNA PCR-free kit following the manufacturer's guidelines then sequenced on the Illumina HiSeq X. Reads were aligned to the hg19 reference sequence (Lander et al. 2001) with the Isaac aligner version 01.15.02.08 (Raczy et al. 2013). Variants were jointly called across the family with GATK HaplotypeCaller v3.5 (McKenna et al. 2010). CNVs were called for each genome by the Reference Coverage Profiles method (Glusman et al. 2015) and merged across the family with SURVIVOR (Jeffares et al. 2017), with maximum allowed distance of 4 kb.

Variant Annotation and Filtering

Variants were annotated using ANNOVAR version 2016-02-04 (Wang et al. 2010) as previously described (Bodian et al. 2017), with the following updates. Transcript models were obtained from the ANNOVAR-provided RefSeq gene (O'Leary et al. 2016) and Ensembl (Yates et al. 2016) databases, version 2017-06-01, and GenCode (Harrow et al. 2012) wgEncodeGencodeCompV24lift37, downloaded from the UCSC genome browser (Kent et al. 2002). Small variants and CNVs predicted to impact the protein sequence of any transcript or that are reported as pathogenic or likely pathogenic in the ClinVar 201705 xml file (Landrum et al. 2016) were identified using SAVANNA as described (Bodian et al. 2017). Briefly, variants were required to have minimum read depth of ≥6, allele balance of

≥0.225, a maximum allele frequency of 0.02 in the gnomAD genomic dataset (Lek et al. 2016), and the data sets provided in the ANNOVAR file popfreq_all_2015_0413, including the 1000 Genomes Project genomes (Auton et al. 2015) and in internal WGS databases. Segregation with phenotype was tested for de novo, autosomal dominant, homozygous recessive, compound heterozygous, X-linked, and imprinted inheritance patterns. Genes linked to the most salient phenotype of the proband, seizures (HP:0001250), as assigned by a clinical geneticist using Phenomizer (Köhler et al. 2009), were retrieved from the HPO database version 01 2016 (Köhler et al. 2017), and variants in those genes were flagged. Sequencing reads supporting the candidate variants were examined visually with IGV (Robinson et al. 2011). Variants resulting from the bioinformatics filtering were then evaluated by clinical genomics experts. Variant cDNA coordinates are reported in the right-shifted HGVS convention (http://varnomen.hgvs.org/).

Sanger Sequencing

Primers were designed, using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/), to cover the *CDKL5* variant identified by WGS (primer sequences available on request). PCR was performed using the AmpliTaq Gold 360 master mix (ThermoFisher Scientific). After Exo/SAP purification (ThermoFisher Scientific), the amplicons were sequenced using BigDye V.3.1 Terminator chemistry (ThermoFisher Scientific). The sequencing products were cleaned up with Optima DTR columns (Edge BioSytems) and separated on an ABI 3730xl genetic analyzer (ThermoFisher Scientific). Data were evaluated using Sequencher V.5.0 software (Gene Codes).

ADDITIONAL INFORMATION

Data Deposition and Access

The confirmed genetic variant was submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) under accession number SCV000599451. Raw data were not deposited because of lack of patient consent but may be available by contacting the authors.

Ethics Statement

The study "The Incidence and Burden of Congenital Anomalies, Genetic Disorders, and Genetic Suspicion in Neonatal and Pediatric Patients" was approved by the Inova Research Center/Inova IRB (IRB# 15-1816) and the Western institutional review board (WIRB #20121680), with full written informed consent obtained from the participants, with the parents providing consent for the proband.

Author Contributions

All authors were involved in the conception or design of the study, or analysis and interpretation of data, and drafting the article or revising it critically for important intellectual content.

REFERENCES

Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR. 2015. A global reference for human genetic variation. *Nature* **526:** 68–74.

Bahi-Buisson N, Bienvenu T. 2012. *CDKL5*-related disorders: from clinical description to molecular genetics. *Mol Syndromol* **2:** 137–152.

Bodian DL, Vilboux T, Hourigan SK, Jenevein CL, Mani H, Kent KC, Khromykh A, Solomon BD, Hauser NS. 2017. Genomic analysis of an infant with intractable diarrhea and dilated cardiomyopathy. *Cold Spring Harb Mol Case Stud* **3:** a002055.

Diebold B, Delépine C, Gataullina S, Delahaye A, Nectoux J, Bienvenu T. 2014. Mutations in the C-terminus of CDKL5: proceed with caution. *Eur J Hum Genet* **22:** 270–272.

Eltze CM, Chong WK, Cox T, Whitney A, Cortina-Borja M, Chin RF, Scott RC, Cross JH. 2013. A population-based study of newly diagnosed epilepsy in infants. *Epilepsia* **54:** 437–445.

Fehr S, Wilson M, Downs J, Williams S, Murgia A, Sartori S, Vecchi M, Ho G, Polli R, Psoni S, et al. 2013. The CDKL5 disorder is an independent clinical entity associated with early-onset encephalopathy. *Eur J Hum Genet* **21:** 266–273.

Fichou Y, Nectoux J, Bahi-Buisson N, Chelly J, Bienvenu T. 2011. An isoform of the severe encephalopathy-related *CDKL5* gene, including a novel exon with extremely high sequence conservation, is specifically expressed in brain. *J Hum Genet* **56:** 52–57.

Galván-Manso M, Campistol J, Conill J, Sanmartí FX. 2005. Analysis of the characteristics of epilepsy in 37 patients with the molecular diagnosis of Angelman syndrome. *Epileptic Disord* **7:** 19–25.

Glusman G, Severson A, Dhankani V, Robinson M, Farrah T, Mauldin DE, Stittrich AB, Ament SA, Roach JC, Brunkow ME, et al. 2015. Identification of copy number variants in whole-genome data using Reference Coverage Profiles. *Front Genet* **6:** 45.

Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22:** 1760–1774.

Hector RD, Dando O, Landsberger N, Kilstrup-Nielsen C, Kind PC, Bailey ME, Cobb SR. 2016. Characterisation of *CDKL5* transcript isoforms in human and mouse. *PLoS One* **11:** e0157758.

Jeffares DC, Jolly C, Hoti M, Speed D, Shaw L, Rallis C, Balloux F, Dessimoz C, Bähler J, Sedlazeck FJ. 2017. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat Commun* **8:** 14061.

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. *Genome Res* **12:** 996–1006.

Köhler S, Schulz MH, Krawitz P, Bauer S, Dölken S, Ott CE, Mundlos C, Horn D, Mundlos S, Robinson PN. 2009. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am J Hum Genet* **85:** 457–464.

Köhler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Aymé S, Baynam G, Bello SM, Boerkoel CF, Boycott KM, et al. 2017. The human phenotype ontology in 2017. *Nucleic Acids Res* **45:** D865–D876.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* **409:** 860–921.

Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, et al. 2016. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* **44:** D862–D868.

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536:** 285–291.

Mangatt M, Wong K, Anderson B, Epstein A, Hodgetts S, Leonard H, Downs J. 2016. Prevalence and onset of comorbidities in the CDKL5 disorder differ from Rett syndrome. *Orphanet J Rare Dis* **11:** 39.

Masliah-Plachon J, Auvin S, Nectoux J, Fichou Y, Chelly J, Bienvenu T. 2010. Somatic mosaicism for a CDKL5 mutation as an epileptic encephalopathy in males. *Am J Med Genet A* **152A:** 2110–2111.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20:** 1297–1303.

Meguro M, Kashiwagi A, Mitsuya K, Nakao M, Kondo I, Saitoh S, Oshimura M. 2001. A novel maternally expressed gene, *ATP10C*, encodes a putative aminophospholipid translocase associated with Angelman syndrome. *Nat Genet* **28:** 19–20.

Mei D, Darra F, Barba C, Marini C, Fontana E, Chiti L, Parrini E, Dalla Bernardina B, Guerrini R. 2014. Optimizing the molecular diagnosis of *CDKL5* gene–related epileptic encephalopathy in boys. *Epilepsia* **55:** 1748–1753.

O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* **44:** D733–D745.

Raczy C, Petrovski R, Saunders CT, Chorny I, Kruglyak S, Margulies EH, Chuang HY, Källberg M, Kumar SA, Liao A, et al. 2013. Isaac: ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. *Bioinformatics* **29:** 2041–2043.

Rademacher N, Hambrock M, Fischer U, Moser B, Ceulemans B, Lieb W, Boor R, Stefanova I, Gillessen-Kaesbach G, Runge C, et al. 2011. Identification of a novel *CDKL5* exon and pathogenic mutations in patients with severe mental retardation, early-onset seizures and Rett-like features. *Neurogenetics* **12:** 165–167.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29:** 24–26.

Sands TT, Choi H. 2017. Genetic testing in pediatric epilepsy. *Curr Neurol Neurosci Rep* **17:** 45.

Sartori S, Di Rosa G, Polli R, Bettella E, Tricomi G, Tortorella G, Murgia A. 2009. A novel *CDKL5* mutation in a 47,XXY boy with the early-onset seizure variant of Rett syndrome. *Am J Med Genet A* **149A:** 232–236.

Trump N, McTague A, Brittain H, Papandreou A, Meyer E, Ngoh A, Palmer R, Morrogh D, Boustred C, Hurst JA, et al. 2016. Improving diagnosis and broadening the phenotypes in early-onset seizure and severe developmental delay disorders through gene panel analysis. *J Med Genet* **53:** 310–317.

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38:** e164.

Weber YG, Biskup S, Helbig KL, Von Spiczak S, Lerche H. 2017. The role of genetic testing in epilepsy diagnosis and management. *Expert Rev Mol Diagn* **17:** 739–750.

Williamson SL, Giudici L, Kilstrup-Nielsen C, Gold W, Pelka GJ, Tam PP, Grimm A, Prodi D, Landsberger N, Christodoulou J. 2012. A novel transcript of cyclin-dependent kinase-like 5 (*CDKL5*) has an alternative C-terminus and is the predominant transcript in brain. *Hum Genet* **131:** 187–200.

Yates A, Akanni W, Amode MR, Barrell D, Billis K, Carvalho-Silva D, Cummins C, Clapham P, Fitzgerald S, Gil L, et al. 2016. Ensembl 2016. *Nucleic Acids Res* **44:** D710–D716.