HHS Public Access

Author manuscript

*Nat Genet*. Author manuscript; available in PMC 2014 January 01.

Published in final edited form as:

*Nat Genet*. 2013 July ; 45(7): 825–830. doi:10.1038/ng.2646.

# Targeted resequencing in epileptic encephalopathies identifies *de novo* mutations in *CHD2* and *SYNGAP1*

**Gemma L. Carvill**[1], **Sinéad B. Heavin**[2], **Simone C. Yendle**[2], **Jacinta M. McMahon**[2], **Brian J. O'Roak**[3], **Joseph Cook**[1], **Adiba Khan**[1], **Michael O Dorschner**[4,5], **Molly Weaver**[4,5], **Sophie Calvert**[6], **Stephen Malone**[6], **Geoffrey Wallace**[6], **Thorsten Stanley**[7], **Ann M. E. Bye**[8], **Andrew Bleasel**[9], **Katherine B. Howell**[10], **Sara Kivity**[11], **Mark T. Mackay**[10,12,13], **Victoria Rodriguez-Casero**[14], **Richard Webster**[15], **Amos Korczyn**[16], **Zaid Afawi**[17], **Nathanel Zelnick**[18], **Tally Lerman-Sagie**[19], **Dorit Lev**[19], **Rikke S. Møller**[20], **Deepak Gill**[15], **Danielle M. Andrade**[21], **Jeremy L. Freeman**[10,12], **Lynette G. Sadleir**[7], **Jay Shendure**[3], **Samuel F. Berkovic**[2], **Ingrid E. Scheffer**[2,10,13,22,*], and **Heather C. Mefford**[1,*]

[1]Division of Genetic Medicine, Department of Pediatrics, University of Washington, Seattle, Washington, 98195, USA [2]Epilepsy Research Centre, Department of Medicine, University of Melbourne, Austin Health, Melbourne, Victoria, Australia [3]Department of Genome Sciences, University of Washington, Seattle, Washington, 98195, USA [4]Department of Psychiatry & Behavioral Sciences, University of Washington, Seattle, Washington, 98195, USA [5]Veteran Affairs Puget Sound Health Care System, Seattle, WA 98108, USA [6]Neurosciences Children's Health Queensland, Royal and Mater Children's Hospitals, Brisbane, Queensland, Australia [7]Department of Paediatrics, School of Medicine and Health Sciences, University of Otago, Wellington, New Zealand [8]Department of Paediatric Neurology, University of New South Wales, Sydney Children's Hospital, Sydney, New South Wales, Australia [9]Westmead Hospital, University of Sydney, Sydney, New South Wales, Australia [10]Department of Neurology, The Royal Children's Hospital, Parkville, Melbourne, Victoria, Australia [11]Epilepsy Unit, Schneider Children's Medical Center of Israel, Petach Tikvah, Israel [12]Critical Care & Neurosciences Theme, Murdoch Childrens Research Institute, Melbourne, Victoria, Australia [13]Department of Paediatrics, The



[*]To whom correspondence should be addressed: Heather C. Mefford, hmefford@uw.edu; Ingrid E. Scheffer, scheffer@unimelb.edu.au.

**Author contribution statement**

GLC, HCM and IES designed the study and wrote the manuscript. HCM and IES supervised the study. GLC constructed libraries, developed the variant calling pipeline (assisted by JC), and analyzed the sequence data. BJO and JS developed the MIPs methodology and analysis pipeline. SBH, SCY, JMM, SC, SM, GW, TS, AMEB, AB, KBH, SK, MTM, VRC, RW, AK, ZA, NZ, TLS, DL, RSM, DG, DMA, JLF, LGS, SFB, IES performed phenotypic analysis. SBH, JMM, SFB and IES critically reviewed the manuscript. GLC and AK performed segregation analysis experiments. MOD and MW performed Illumina Hiseq sequencing.

**Resources**

Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [10/2012 accessed]
Genome Analysis Toolkit GATK (version 2.2) (http://www.broadinstitute.org/gatk/)
Seattle seq (version 134) (http://snp.gs.washington.edu/SeattleSeqAnnotation134/)
BWA; Burrows-Wheeler Aligner (version 0.5.9) (http://bio-bwa.sourceforge.net/)
Picard tools (version 1.82) (http://picard.sourceforge.net/)
Polyphen2 (URL: http://genetics.bwh.harvard.edu/pph2/)
SIFT, Sorting Intolerant from Tolerant, (URL: http://sift.bii.a-star.edu.sg/)

University of Melbourne, The Royal Children's Hospital, Melbourne, Victoria, Australia [14]Paediatric Neurology, Monash Medical Centre, Melbourne, Victoria, Australia [15]TY Nelson Department of Neurology, The Children's Hospital at Westmead, Sydney, New South Wales, Australia [16]Department of Neurology, Tel-Aviv University, Tel-Aviv, Israel [17]Tel-Aviv University Medical School, Tel-Aviv, Israel [18]Department of Pediatrics, Carmel Medical Center, Technion Faculty of Medicine, Haifa, Israel [19]Metabolic-Neurogenetic Service, Wolfson Medical Center, Holon, and Sackler School of Medicine, Tel-Aviv University, Tel-Aviv, Israel [20]Danish Epilepsy Centre, Dianalund, Denmark [21]Division of Neurology, Department of Medicine, University of Toronto, Toronto Western Hospital, Krembil Neurosciences Program. Toronto, Canada [22]Florey Institute, Melbourne, Victoria, Australia

## Introductory

Epileptic encephalopathies are a devastating group of epilepsies with a poor prognosis, for which the majority have unknown etiology. We perform targeted massively parallel resequencing of 19 known and 46 candidate epileptic encephalopathy genes in 500 patients to identify novel genes and investigate the phenotypic spectrum of known genes. Overall, we identify pathogenic mutations in 10% of our cohort. Six of the 46 candidate genes had one or more pathogenic variants, collectively accounting for 3% of our cohort. We show that *de novo CHD2* and *SYNGAP1* mutations are novel causes of epileptic encephalopathies, accounting for 1.2% and 1% of cases respectively. We also further expand the phenotypic spectrum for *SCN1A*, *SCN2A*, and *SCN8A* mutations. To our knowledge, this is the largest cohort of patients with epileptic encephalopathies to undergo targeted resequencing. Implementation of this rapid and efficient method will change diagnosis and understanding of the molecular etiologies of these disorders.

Epilepsy is one of the most common neurological disorders with a lifetime incidence of 3%. Epileptic encephalopathies are a devastating group of epilepsies characterized by refractory seizures and cognitive arrest or regression associated with ongoing epileptic activity, and typically carry a poor prognosis[1]. *De novo* mutations in several known genes are responsible for some epileptic encephalopathies[2]. Furthermore, we and others have shown that rare, *de novo* copy number variants (CNVs) account for up to ~8% of cases[3, 4]. Despite this recent progress, making a genetic diagnosis in a patient can be challenging as there is both genetic heterogeneity for a given epilepsy syndrome and phenotypic heterogeneity for a specific gene.

The full phenotypic spectrum associated with mutations in known epileptic encephalopathy genes is not known. Very few studies have investigated the role of any given gene across a wide spectrum of epileptic encephalopathy syndromes. This makes serial gene testing in the clinical setting an inefficient and expensive process, after which the vast majority of cases remain unexplained. Furthermore, it is clear that discovery of additional genes that cause epileptic encephalopathies is needed to facilitate genetic diagnosis. Here, we take advantage of a high-throughput targeted sequencing approach to perform comprehensive sequence analysis of 65 genes (19 known genes and 46 candidate genes) (Supplementary Fig. 1) in

500 patients with a range of epileptic encephalopathy phenotypes (Table 1). Candidate genes were selected from epilepsy-associated CNVs (n=33) or because mutations cause associated neurodevelopmental disorders or other epilepsy syndromes (n=13). Using this approach, we (i) identify novel epileptic encephalopathy genes and (ii) delineate the phenotypic spectrum and mutation frequency for both known and novel epileptic encephalopathy genes.

Overall, 91% of the target (65 genes) was sequenced at >25X coverage, required for accurate variant calling (Supplementary Fig. 2). We achieved 91% sensitivity across 685 variants (161 loci) from 12 samples that had previously undergone exome sequencing and 100% sensitivity for 24 known variants in previously tested patients; these patients were not included in the discovery cohort.

We detected one or more pathogenic or likely pathogenic mutations in six of our 46 candidate genes, with multiple individuals carrying mutations in either of the two novel epileptic encephalopathy genes, *CHD2* (NM_001271.3, NP_001262.3) and *SYNGAP1* (NM_006772.2, NP_006763.2) (Table 1, 2, Fig. 1).

Remarkably, we detected six *de novo* variants in the candidate gene, *CHD2* (Fig. 1,2), selected from within the critical interval of 15q26.1 deletions detected in patients with a range of epileptic encephalopathies (Supplementary Fig. 3)[5, 6]. Four mutations lead to premature truncation of CHD2 (Table 2). Two *de novo* missense variants disrupt highly conserved residues within the SNF2-related helicase/ATPase domain (p.Trp548Arg and p.Leu823Pro), and are predicted to be damaging by both PolyPhen2 and SIFT. *CHD2* codes for a member of the chromodomain helicase DNA-binding family of proteins and is characterized by the presence of chromatin remodeling, chromo (chromatin organization modifier) and SNF2-related helicase/ATPase domains. These domains suggest function of this protein as a chromatin remodeler[7]. While functional studies in CHD2 are limited, studies of another CHD protein family member, CHD7, have shown that the helicase domain is responsible for ATP-dependent nucleosome remodeling, an integral process in target gene regulation. Furthermore, *in vivo* studies of human *CHD7* mutations within the helicase domain, which cause CHARGE syndrome, resulted in decreased remodeling ability[8]. These results suggest that the two *de novo* missense mutations described here may disrupt CHD2 function in a similar manner, while truncating mutations likely result in haploinsufficiency.

The six patients with *CHD2* mutations had distinctive features with a median seizure onset of 18 months (range 1–3 years, Table 2): myoclonic seizures in all, photosensitivity in three and all had ID, ranging from moderate to severe. A *de novo CHD2* frameshift mutation was reported in a proband with ID and absence seizures[9] and a *de novo* missense mutation in an individual with autism spectrum disorder (ASD)[10]. These results suggest that mutations in *CHD2* contribute to a broad spectrum of neurodevelopmental disorders. Notably, recent studies implicate *de novo* mutations in *CHD8* in patients with ASD[11]. Interestingly, three genes of the chromodomain family (*CHD2*, *CHD7*, *CHD8*) have now been implicated in disorders that impact the neurodevelopmental system. Further studies of this nine-member gene family will determine the role of each across the spectrum of neurodevelopmental disorders, and provide exciting new avenues of research.

We identified nine pathogenic or likely pathogenic variants in four of the 13 'epilepsy-associated' genes (Fig. 1). We found five truncating variants in *SYNGAP1* (Fig. 2). Patients with *SYNGAP1* mutations had median seizure onset of 14 months (mean 14 months, range 6 months to 3 years) (Table 2). They had multiple seizure types, early developmental delay and subsequent regression. Outcome was poor with moderate to severe ID. *SYNGAP1* mutations have been associated with ID and, although most patients have epilepsy, seizures are typically well controlled[9, 12–18]. Our study represents the first cases of epileptic encephalopathies with *SYNGAP1* mutations. These observations suggest that epilepsy is a core feature of both static and progressive encephalopathies associated with *SYNGAP1* mutations, and carry important implications for diagnostic testing.

Variants were identified in three additional 'epilepsy associated genes'. There were two *de novo* variants in *MEF2C* (NM_002397.4, NP_002388.2), a missense variant and a stop-loss variant (p.*464SerExt*?). Furthermore, we found *de novo* pathogenic variants in *MBD5* (NM_018328.4, NP_060798.2) (Thr157Glnfs*4) and *GABRG2* (NM_000816.3, NP_000807.2)(p.Arg323Gln) (Table 2).

We detected a premature truncation mutation (p.Tyr805*) in the CNV candidate gene, *HNRNPU* (NM_031844.2, NP_114032.2). The p.Tyr805* change arose as a result of two consecutive single nucleotide changes c.471T>C and c.472A>T (Supplementary Fig. 4) that occur two amino acids upstream of the termination codon. Neither variant was maternally inherited; paternal DNA was not available. A recent report identified *HNRNPU* as a candidate for the ID and seizure phenotypes of probands with 1q44 microdeletions[19]. In addition, a *de novo* splice-site variant was identified in a proband with a complex neurodevelopmental phenotype including epilepsy[20]. Collectively, these data suggest that haploinsufficiency of *HNRNPU* is associated with epileptic encephalopathy as well as ID, though further phenotype-genotype correlation will improve our understanding of the *HNRNPU* phenotypic spectrum.

We identified 32 variants fulfilling our criteria for pathogenicity and an additional four variants that are likely pathogenic in ten of 19 known epileptic encephalopathy genes (Fig. 1, Table 1, Table 3). We identified multiple patients with mutations in *STXBP1*, *CDKL5*, *SCN1A*, *SCN2A*, *PCDH19* and *KCNQ2*, accounting for 69% (36/52) of all mutation-positive individuals in our cohort. We detected an additional 16 rare variants in six of these 19 known genes for which we were unable to conduct segregation analysis; it is probable that a number of these variants are also pathogenic (Supplementary Table 1).

The phenotypes identified in patients with mutations in known genes are provided (Table 3), and for some we expand the known phenotypic spectrum. For example, we identified a homozygous recessive missense mutation in *PNKP* in a single proband with unclassified epileptic encephalopathy. *PNKP* mutations are associated with early infantile epileptic encephalopathy comprising microcephaly, early-onset intractable seizures and developmental delay[21]. By contrast, our patient did not have microcephaly (head circumference 50$^{th}$ centile) or developmental delay but had normal cognition despite refractory epilepsy with multiple seizure types. Also, three patients with *SCN1A* mutations presented with an epilepsy-aphasia phenotype, of which two also had FS+. *SCN1A*

mutations are well known to be associated with genetic epilepsy with febrile seizures plus (GEFS+) but have not previously been reported with epilepsy-aphasia syndromes[22, 23]. It is possible that the *SCN1A* mutation is not responsible for the epilepsy-aphasia syndrome but equally it could be a modifier predisposing the individual to this group of epileptic encephalopathies. Further work is warranted to clarify this association, perhaps most effectively with exome-sequencing in these patients.

We detected five variants in *SCN2A*, which encodes the α2 subunit of the voltage gated sodium channel. To date, the majority of *SCN2A* mutations have been associated with the self-limited autosomal dominant syndrome of benign familial neonatal-infantile seizures (BFNIS)[24]. Previously, only three *de novo* variants have been reported in patients with epileptic encephalopathies[25, 26]. Interestingly our five cases show similar variability in the range of onset seen in BFNIS with three beginning in the neonatal period (11 hours to 2 days) and two in infancy (6 weeks, 13 months). Two had relatively early offset of seizures at 5 weeks and 7 months. The refractory nature of seizures did not correlate with intellectual outcome, which ranged from mild (2) to severe (3) intellectual disability. We conclude that *SCN2A* is an important contributor to the overall burden of epileptic encephalopathies, accounting for 1% of cases.

We also identified a pathogenic missense mutation (p.Leu1290Val) in *SCN8A*. To date, only a single *de novo SCN8A* mutation (p.Asn1768Asp) has been described in a proband with severe epileptic encephalopathy and sudden unexplained death in epilepsy[27]. Here we describe a second patient presenting with an epileptic encephalopathy beginning at 18 months. Interestingly, this variant was paternally inherited, though the father was shown to have somatic mosaicism (13% mutant allele) supporting its pathogenic effect as seen in other genetic encephalopathies with parental mosaicism[28].

The findings in this large series of patients with hitherto unsolved epileptic encephalopathies allows us to begin to frame the overall genetic architecture of this group of disorders. We identified pathogenic or likely pathogenic mutations in 10% of our cohort, with mutations in 16 genes. However, this mutation rate is likely to be an underestimation of the true contribution of each gene to the overall burden of epileptic encephalopathies. Our cohort excluded patients with previously identified mutations, and we were unable to conduct segregation analysis for a subset of variants we identified, some of which are likely to be pathogenic. Furthermore, as larger numbers of patients with mutations of specific genes are identified, distinctive epileptic encephalopathy phenotypes are likely to emerge. Taken together, with up to 8% rare CNVs in epileptic encephalopathy patients in an earlier analysis of a subset of this series[3], we can now collectively ascribe causality for ~18% of all epileptic encephalopathies of unknown cause.

The genetic heterogeneity of epileptic encephalopathies is considerable; likely pathogenic variants were found in nine known or novel genes (see Fig. 2). Even the most commonly mutated genes in our study each account for only up to 1.6% of cases. Notably, we elucidate new genes found to be commonly mutated in epileptic encephalopathies, with *CHD2*, *SYNGAP1* and *SCN2A* accounting for 1–1.2% of cases each, a frequency similar to that of mutations in *SCN1A*, *STXBP1* and *CDKL5* in our cohort. However, no mutations were seen

in nine other known genes (*ARX*, *FOXG1*, *KCNT1*, *MECP2*, *PLCB1*, *SLC25A22*, *SLC2A1*, *SPTAN1*, *ARHGEF9*) in 500 patients. These results suggest that pathogenic mutations in these genes, while important, are rare causes of epileptic encephalopathies (<0.2% each in our cohort), or cause only very distinct syndromes that were not prevalent in our cohort. These findings support a clinical approach to genetic diagnosis that employs large gene panels or whole exome sequencing, as it will remain difficult and expensive to determine *a priori* the causative gene in a given patient.

Notably, mutations in *SYNGAP1* and *CHD2* have now been described in probands with epileptic encephalopathy, ID and ASD phenotypes, highlighting the shared genetic basis of neurodevelopmental disorders. Unbiased approaches such as exome or whole genome sequencing provide an avenue to gene discovery, but large cohorts will be required to identify two or more patients with *de novo* mutations in the same gene[9, 18]. Our results show the power of targeted resequencing to screen large numbers of patients in a high-throughput and cost-effective manner. This approach is critical to identify additional patients with mutations in genes where a single *de novo* mutation is identified by exome sequencing approaches, to determine overall mutation frequency in a given phenotype, and to describe genotype-phenotype correlations. Applying this approach across various neurodevelopmental disorders will identify additional mutation positive patients for a specific gene and enhance our understanding of disease mechanisms.

## ONLINE METHODS

### Patients

This study was approved by the Human Research Ethics Committees of Austin Health and the University of Washington. Probands with epileptic encephalopathies were recruited from the epilepsy clinic at Austin Health, the practices of the investigators and by referral for epilepsy genetics research from around Australia and internationally after informed consent. The cohort consisted of 500 patients with a diverse range of epileptic encephalopathy phenotypes. An epileptic encephalopathy was defined as refractory seizures and cognitive slowing or regression associated with frequent, ongoing epileptiform activity[1]. Detailed epilepsy and medical histories were obtained together with the results of investigations including EEG and MRI studies. Epilepsy syndromes were classified according to the Organization of the International League Against Epilepsy Commission on Classification[1] (Table 1). Some patients had already undergone mutation screening for specific epileptic encephalopathy genes; none had been screened for all known genes. Patients with a previously identified disease causing mutation were excluded from this study. Some patients with known mutations were included as mutation positive controls but were not included in the 500 cases in the discovery cohort. Furthermore, 369/500 patients had been previously screened for pathogenic CNVs, either in the research setting (n=257)[3], or by clinical testing (n=112). Probands with pathogenic CNVs were not included in this study.

### Gene Selection

We selected 65 genes for sequence analysis (Fig. 1). The “known” gene group included genes in which mutations are known to cause one or more epileptic encephalopathy

syndromes (n=19)[2, 21, 25, 27, 28, 32–34]. We also selected two sets of candidate genes for epileptic encephalopathies. The first set includes 13 'epilepsy-associated genes', more commonly implicated in patients with non-epileptic encephalopathy forms of epilepsy (*CHRNA7, KCNQ3, GABRD, GABRG2, PRICKLE1, CACNB4, SCN1B*) or a related neurodevelopmental disorder with epilepsy as a comorbid feature (*GRIN2B, MBD5, MEF2C, SYNGAP1, SYN1, ATP1A2*)[35–39]. None are an established cause of epileptic encephalopathies. We also selected 33 candidate genes, primarily from epilepsy-associated CNVs, either from published cohorts, case reports or unpublished data[3, 4, 40] (see Supplementary Note and Supplementary Table 2 for candidate gene selection).

## Controls

Sixteen samples that had been previously subject to exome sequencing were included in all analyses and used to assess the sensitivity of variant calling. We also included 24 probands with a known variant in a known or candidate gene to further validate our approach.

## Target capture and sequencing

We used Molecular Inversion Probes (MIPs) to capture all exon and intron/exon boundaries (5bp flanking) of target genes (Refseq, hg19 build) (Supplementary Table 3). Detailed methodology is described elsewhere[11]. Briefly, pooled MIPs were used to capture target exons from 100ng of each proband's DNA. PCR was performed using universal primers with the introduction of unique eight-base barcodes on the tagged reverse primer. Pooled libraries were subject to massively parallel sequencing using a 101 paired-end protocol on a Hiseq. The libraries were prepared and sequenced in two batches, comprising a total of 30 (target 1) and 35 (target 2) genes.

## Data analysis and variant calling

Raw read data processing and mapping with BWA (See URLs) was performed as described[11]. Single nucleotide variant (SNV) and indel calling and filtering was performed using the Genome Analysis Tool Kit (GATK) (see URLs). Variants that did not adhere to the following criteria were excluded from further analysis: allele balance >0.70, QUAL<30, QD<5, coverage<25X, clustered variants (window size-10). Variants were annotated with Seattle seq (see URLS) and the ESP6500 dataset (see URLs) was used to assess variant frequency in the control population. For dominant (or *de novo*) models we considered only variants not present in this control sample set. For recessive candidates, we considered variants with a frequency in controls of <1%. Only non-synonymous, splice-site or frameshift ('damaging') variants were assessed further. The GATK Depth of Coverage tool was used to calculate overall depth of coverage for each sample at a threshold of 25X, as well as the mean percentage (across all samples) of bases covered >25X for each gene.

## Rare variant segregation analysis

Where family members were available, segregation analysis was carried out for all rare (not present in ESP6500 controls), possibly damaging (non-synonymous, essential splice-site or frameshift) variants for all 65 target genes. This analysis was performed using a 'MIP-pick' strategy. We selected and re-pooled only the MIPs that captured the genomic sequence

harboring the rare variant of interest and performed target enrichment PCR and sequencing as above for all relevant probands and family members. This approach allowed us to sequence variants at very high depth and detect somatic mosaicism in parents.

### Criteria for pathogenicity of rare variants

For those rare, possibly damaging variants where segregation analysis could be performed, we required the variant to meet one of the following criteria to constitute a novel pathogenic variant. Pathogenic variants: (i) arose *de novo*, (ii) segregated with the disorder, (iii) were inherited from a parent with somatic mosaicism, or (iv) adhered to a recessive, X-linked or parent-of-origin mode of inheritance, where applicable (Supplementary Fig. 1).

In certain instances we were unable to determine the inheritance of a rare variant due to the unavailability of DNA from one or more parent. It is likely that a subset of these variants also cause disease, though here we report only those variants that are likely to lead to protein truncation (i.e. splice-site, nonsense, frameshift, stop-loss) as being 'likely pathogenic'. Additionally, two missense mutations in known genes (*STXBP1, SCN2A*) were interpreted to be 'likely pathogenic' based on the high incidence of pathogenic missense mutations in these genes, which was further supported by the available parent not carrying the variant. We performed microsatellite analysis using the PowerPlex S5 system [Promega] in all parents of probands with a *de novo* mutation to confirm maternity and paternity.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank the patients and their families for participating in our research. H.C.M. is supported by a grant from the NIH (NINDS 1R01NS069605) and is a recipient of a Burroughs Wellcome Fund Career Award for Medical Scientists. This work was supported by the National Health and Medical Research Council of Australia (Program Grant 628952 to S.F.B., I.E.S., Practitioner Fellowship 1006110 to I.E.S and the Health Research Council of New Zealand project grant to L.G.S.)

## References

1. Berg AT, et al. Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology 2005–2009. Epilepsia. 2010; 51:676–685. [PubMed: 20196795]
2. Kamien BA, Cardamone M, Lawson JA, Sachdev R. A genetic diagnostic approach to infantile epileptic encephalopathies. J. Clin. Neurosci. 2012; 19:934–941. [PubMed: 22617547]
3. Mefford HC, et al. Rare copy number variants are an important cause of epileptic encephalopathies. Ann. Neurol. 2011; 70:974–985. [PubMed: 22190369]
4. Heinzen EL, et al. Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. Am. J Hum. Genet. 2010; 86:707–718. [PubMed: 20398883]
5. Capelli LP, et al. Deletion of the RMGA and CHD2 genes in a child with epilepsy and mental deficiency. Eur. J. Med. Genet. 2012; 55:132–134. [PubMed: 22178256]
6. Dhamija R, et al. Microdeletion of chromosome 15q26.1 in a child with intractable generalized epilepsy. Pediatr. Neurol. 2011; 45:60–62. [PubMed: 21723464]
7. Marfella CG, Imbalzano AN. The Chd family of chromatin remodelers. Mutat. Res. 2007; 618:30–40. [PubMed: 17350655]

8. Bouazoune K, Kingston RE. Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. Proc. Natl. Acad. Sci. U.S.A. 2012; 109:19238–19243. [PubMed: 23134727]
9. Rauch A, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet. 2012
10. Neale BM, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature. 2012; 485:242–245. [PubMed: 22495311]
11. O'Roak BJ, et al. Multiplex Targeted Sequencing Identifies Recurrently Mutated Genes in Autism Spectrum Disorders. 2012 Published online November 15, 2012.
12. Hamdan FF, et al. Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. N. Engl. J. Med. 2009; 360:599–605. [PubMed: 19196676]
13. Hamdan FF, Gauthier J, Rouleau GA, Michaud JL. De novo mutations in SYNGAP1 associated with non-syndromic mental retardation. Med. Sci. (Paris). 2010; 26:133–135. [PubMed: 20188038]
14. Hamdan FF, et al. De novo SYNGAP1 mutations in nonsyndromic intellectual disability and autism. Biol. Psychiatry. 2011; 69:898–901. [PubMed: 21237447]
15. Hamdan FF, et al. Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. Am. J. Hum. Genet. 2011; 88:306–316. [PubMed: 21376300]
16. Berryer MH, et al. Mutations in SYNGAP1 Cause Intellectual Disability, Autism and a Specific form of Epilepsy by Inducing Haploinsufficiency. Hum. Mutat. 2012
17. Vissers LE, et al. A de novo paradigm for mental retardation. Nat. Genet. 2010; 42:1109–1112. [PubMed: 21076407]
18. de Ligt J, et al. Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability. N. Engl. Med. 2012
19. Thierry G, et al. Molecular characterization of 1q44 microdeletion in 11 patients reveals three candidate genes for intellectual disability and seizures. Am. J. Med. Genet. A. 2012; 158A:1633–1640. [PubMed: 22678713]
20. Need AC, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. J. Med. Genet. 2012; 49:353–361. [PubMed: 22581936]
21. Shen J, et al. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat. Genet. 2010; 42:245–249. [PubMed: 20118933]
22. Tsai M, et al. Clinical genetic study of the epilepsy-aphasia spectrum. Epilepsia. 2012
23. Escayg A, et al. Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. Nat Genet. 2000; 24:343–345. [PubMed: 10742094]
24. Oliva M, Berkovic SF, Petrou S. Sodium channels and the neurobiology of epilepsy. Epilepsia. 2012; 53:1849–1859. [PubMed: 22905747]
25. Ogiwara I, et al. De novo mutations of voltage-gated sodium channel alphaII gene SCN2A in intractable epilepsies. Neurology. 2009; 73:1046–1053. [PubMed: 19786696]
26. Kamiya K, et al. A nonsense mutation of the sodium channel gene SCN2A in a patient with intractable epilepsy and mental decline. J. Neurosci. 2004; 24:2690–2698. [PubMed: 15028761]
27. Veeramah KR, et al. De novo pathogenic SCN8A mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. Am. J. Hum. Genet. 2012; 90:502–510. [PubMed: 22365152]
28. Weckhuysen S, et al. KCNQ2 encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. Ann. Neurol. 2012; 71:15–25. [PubMed: 22275249]
29. Fujiwara T, et al. Mutations of sodium channel alpha subunit type 1 (SCN1A) in intractable childhood epilepsies with frequent generalized tonic-clonic seizures. Brain. 2003; 126:531–546. [PubMed: 12566275]
30. Camprubi C, et al. Novel UBE3A mutations causing Angelman syndrome: different parental origin for single nucleotide changes and multiple nucleotide deletions or insertions. Am J. Med. Genet. A. 2009; 149A:343–348. [PubMed: 19213023]

31. Mills PB, et al. Neonatal epileptic encephalopathy caused by mutations in the PNPO gene encoding pyridox(am)ine 5'-phosphate oxidase. Hum. Mol. Genet. 2005; 14:1077–1086. [PubMed: 15772097]
32. Heron SE, et al. Missense mutations in the sodium-gated potassium channel gene KCNT1 cause severe autosomal dominant nocturnal frontal lobe epilepsy. Nat. Genet. 2012; 44:1188–1190. [PubMed: 23086396]
33. Marco EJ, et al. ARHGEF9 disruption in a female patient is associated with X linked mental retardation and sensory hyperarousal. BMJ Case Rep. 2009; 2009 Epub 2009 Jul 2.
34. Mastrangelo M, Leuzzi V. Genes of early-onset epileptic encephalopathies: from genotype to phenotype. Pediatr. Neurol. 2012; 46:24–31. [PubMed: 22196487]
35. Tao H, et al. Mutations in prickle orthologs cause seizures in flies, mice, and humans. Am. J. Hum. Genet. 2011; 88:138–149. [PubMed: 21276947]
36. Endele S, et al. Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. Nat. Genet. 2010; 42:1021–1026. [PubMed: 20890276]
37. Hoppman-Chaney N, Wain K, Seger P, Superneau D, Hodge J. Identification of single gene deletions at 15q13.3: further evidence that CHRNA7 causes the 15q13.3 microdeletion syndrome phenotype. Clin. Genet. 2012
38. Talkowski ME, et al. Assessment of 2q23.1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. Am. J. Hum. Genet. 2011; 89:551–563. [PubMed: 21981781]
39. Zweier M, et al. Mutations in MEF2C from the 5q14.3q15 microdeletion syndrome region are a frequent cause of severe mental retardation and diminish MECP2 and CDKL5 expression. Hum. Mutat. 2010; 31:722–733. [PubMed: 20513142]
40. Mefford HC, et al. Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS Genet. 2010; 6:e1000962. [PubMed: 20502679]

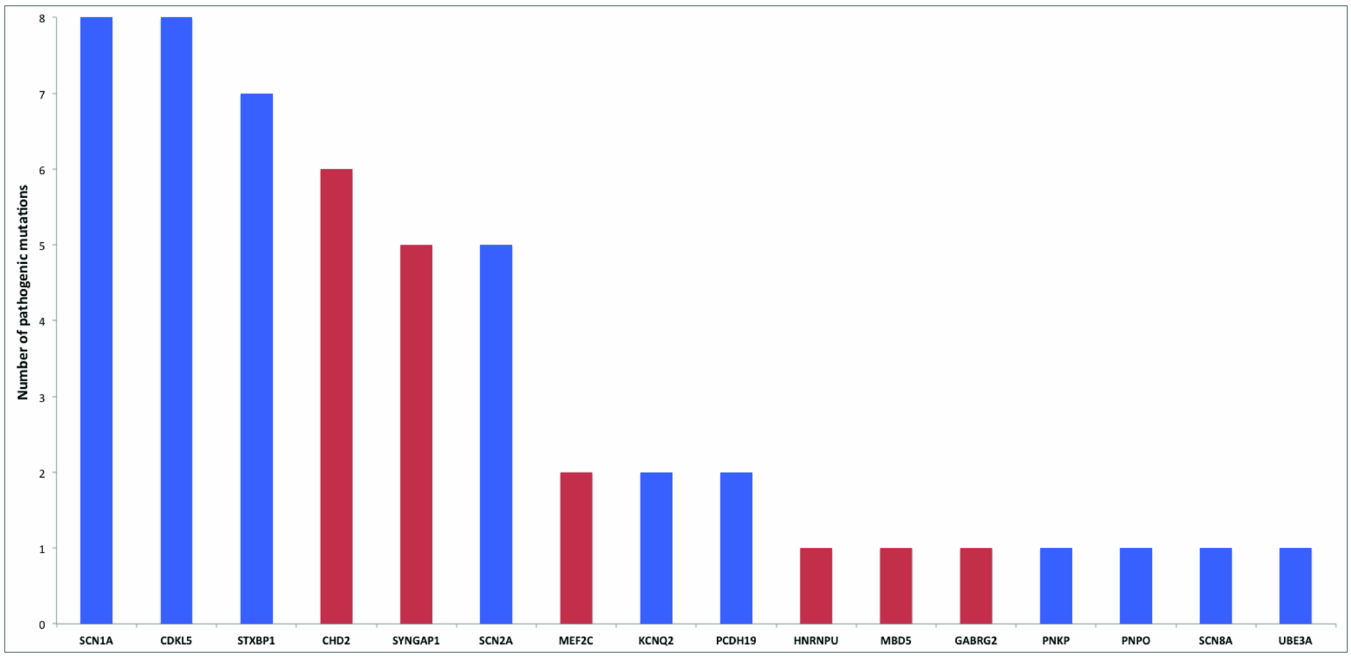


**Figure 1.**

Pathogenic and likely pathogenic mutations identified in 500 patients with epileptic encephalopathies in novel genes (red) and known genes (blue)

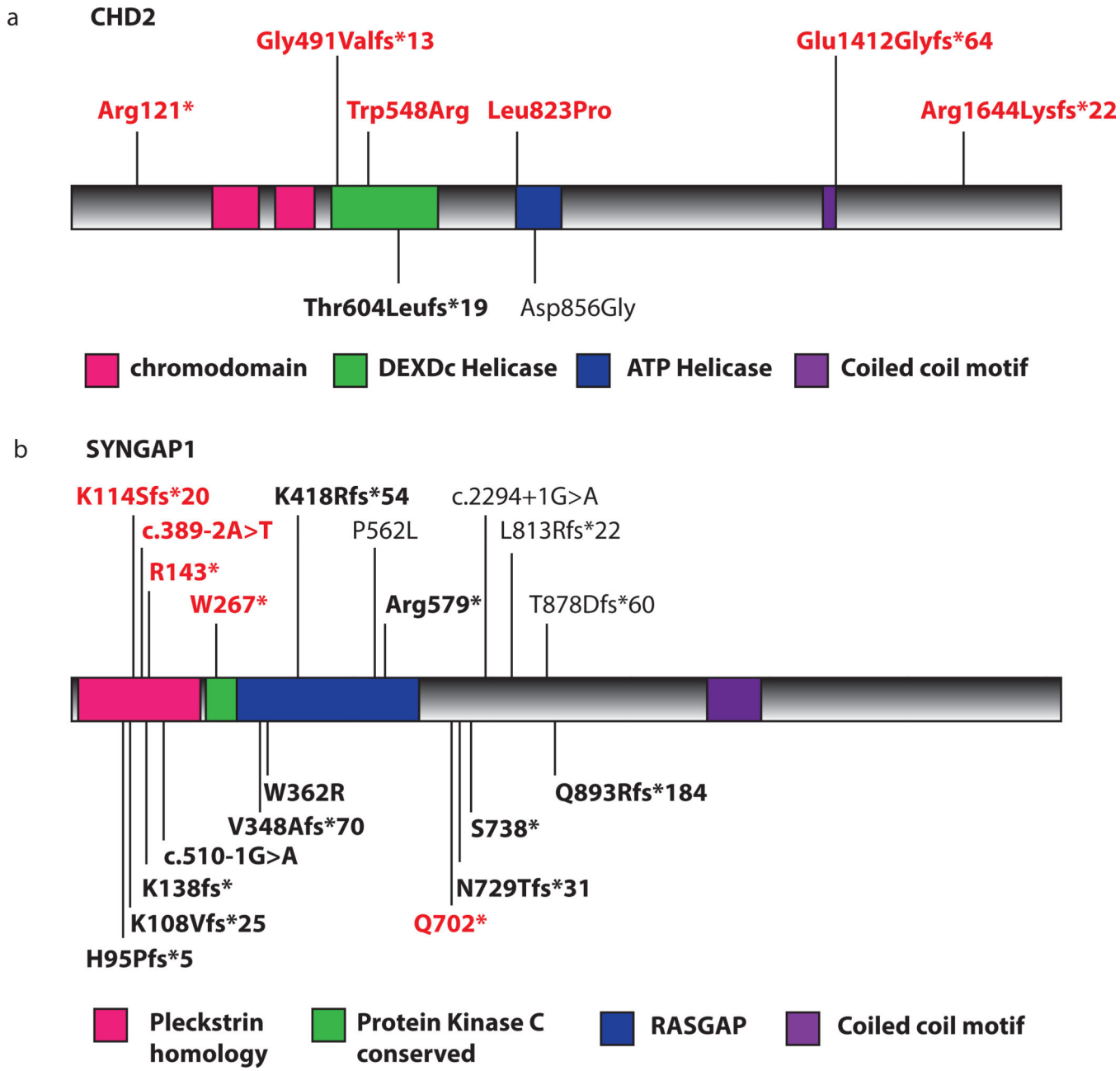


**Figure 2.**

*De novo* mutations in novel epileptic encephalopathy genes a) *CHD2* and b) *SYNGAP1*. Mutations shown in red were identified in this study. Black entries denote previously reported variants; for CHD2, in ID (Thr604Leufs*19)[9] and autism (Asp856Gly)[10] and for SYNGAP1 in ID and/or autism[9, 14–18]. Bold entries indicate pathogenic variants found in patients with epilepsy. No evident genotype-phenotype correlations exist for mutations in either CHD2 or SYNGAP1. For both genes, truncating and missense mutations occur in all three phenotypes (ID/epileptic encephalopathy /autism) without phenotype-specific

intragenic localization. This suggests that alternative neurobiological conditions and mechanisms, genetic or otherwise, underlie this heterogeneity.

## Table 1

Epileptic encephalopathy cohort screened for mutations in 65 novel and known genes

| Syndrome | N | Pathogenic or likely pathogenic variant (% of syndrome) |
|---|---|---|
| ABPE | 6 | 0 |
| Dravet | 19 | 4 (21%)† |
| ECSWS | 10 | 0 |
| Epileptic encephalopathy not otherwise specified | 173 | 22 (13%) |
| EME | 5 | 1 (20%) |
| EOEE | 39 | 8(21%) |
| Epilepsy-Aphasia | 27 | 3 (11%) |
| FIRES | 12 | 0 |
| IS | 81 | 4(5%) |
| LKS | 3 | 0 |
| LGS | 40 | 5 (13%) |
| MAE | 81 | 3 (4%) |
| Ohtahara | 4 | 2 (50%) |
| **TOTAL** | **500** | **52 (10%)** |

ABPE = Atypical Benign Partial Epilepsy; ECSWS, = Epileptic encephalopathy with Continuous Spike-and-Wave during Sleep; EME = Early Myoclonic Encephalopathy, EOEE = Early Onset Epileptic Encephalopathy; FIRES = Febrile Infection-Related Epilepsy Syndrome; IS = Infantile Spasms; LKS = Landau-Kleffner Syndrome; LGS = Lennox-Gastaut Syndrome; MAE = Myoclonic Atonic Epilepsy

† Note that in 4 of the Dravet syndrome cases included here, *SCN1A* testing had not yet been undertaken.

**Table 2**

Pathogenic and likely pathogenic variants in novel epileptic encephalopathy genes

| Proband (gender, study age) | Gene | Protein change (Polyphen, SIFT)[&] | Diagnosis | Seizures[$] (age of onset) | EEG | Development prior to seizure onset | Cognitive outcome (regression) |
|---|---|---|---|---|---|---|---|
| **Pathogenic variants (*de novo*)** | | | | | | | |
| T38 (M, 17 yr) | *CHD2* | Glu1412Glyfs*64 | MAE | At (12mth) FS, Ab, MJ-At, MJ, TC | 3.8 Hz GSW | mild delay, behavioral problems | moderate ID, ASD (no) |
| T18697 (F, 12 yr) | *CHD2* | Arg121* | EE | MJ(12mth) NCS, T, TC, MAb | GPSW, MFD, GPFA, SSW | Normal | severe ID (yes) |
| T2608 (F, 29 yr) | *CHD2* | Gly491Valfs*13 | LGS | aAb (12mth) At, MJ, NCS, SE, T, TC | SSW, MFD, DS, GPSW, PPR triggered MJ | Delayed | severe ID (yes) |
| T20240 (M, 12 yr) | *CHD2* | Arg1644Lysfs*22 | MAE | At (2y) MJ, SE, TC | DS, GPSW, 2.5 Hz GSW | Normal | severe ID (yes) T17756 |
| T17756 (M, 15 yr) | *CHD2* | Trp548Arg (1, 0) | EE | TC (3y) FDS, H, MJ | GSW, MFD, DS | Delayed | Moderate ID (yes) |
| T18431 (M, 2.5 yr) | *CHD2* | Leu823Pro (0.999, 0) | EE | FDS, MJ (2.5y) MJ-Ab, T | GSW, GPSW, MFD | Delayed | Severe ID, ASD (yes) |
| T15923 (F, 26 yr) | *SYNGAP1* | Trp267* | EE | aAb(3y) At, Aura, FDS, MJ | SSW, MFD | Delayed | Severe ID, ASD (yes) |
| T2528 (M, 26 yr) | *SYNGAP1* | Gln702* | EE | FS (18mth) Ab, Aura, FDS, MJ, NCS, TC | SSW, bi-occipital ED, DS | Delayed | Moderate ID (yes) |
| T1898 (F, 20 yr) | *MBD5* | Thr157Glnfs*4 | EE | TC (6mth) Ab, FDS, Focal,,T | GPSW, MFD, DS | Delayed | Severe ID, ASD (no) |
| T20719 (M, 2.5 yr) | *GABRG2* | Arg323Gln (0.998,0) | MAE | FS (8mth) Ab. At, MJ, TC | Normal | Normal | Normal (yes) |
| T23549 (F, 3.5 yr) | *MEF2C* | Cys39Arg (0.998, 0) | EE | FS (13mth) Ab, SE, TC | MFD, DS | Delayed | Severe ID, ASD (yes) |
| T18044[+] (M, 4 yr) | *MEF2C* | *464Sext*? | EE | H (4 mth) Ab, At, focal, IS, MJ, TC | Mod Hyps, MFD | Normal | ID (no) |
| **Likely pathogenic** | | | | | | | |
| T19988[+] (M, 18 yr) | *SYNGAP1* | Lys108Valfs*25 | EE | Unk (in foster care) FDS | MFD, DS | Unk | Moderate ID ASD (unk) |
| T15924[#] (M, 11 yr) | *SYNGAP1* | Unk (c.389-2A>T) | EE | Ab (6mth) TC | GSW, GPSW, MFD | Delayed | Severe ID, ASD (yes) |
| T22387[+] (F, 7 yr) | *SYNGAP1* | Arg143* | EE | Ab(10mth) MJ | GSW | Delayed | Severe ID, ASD (yes) |
| T162[#] (M, 33 yrs) | *HNRNPU* | Tyr805* | LGS | At (2 y) aAb, MJ, NCS, T, TC | GSW, GPSW, DS, SSW, GPFA | Delayed | Severe ID (yes) |

[&]Polyphen and SIFT scores given for missense variants only.

[$]Initial seizure listed first (age of onset, mth – months, y - years), subsequent seizure types. M- male F-female.

[+]Parents unavailable,

[#]father unavailable.

Accession numbers: *CHD2*, NM_001271.3, NP_001262.3; *SYNGAP1*, NM_006772.2, NP_006763.2; *MBD5*, NM_018328.4, NP_060798.2; *GABRG2*, NM_000816.3, NP_000807.2; *MEF2C*, NM_002397.4, NP_002388.2; *HNRNPU*, NM_031844.2, NP_114032.2.

EE = Epileptic encephalopathy; LGS = Lennox Gastaut syndrome; MAE = Myoclonic atonic epilepsy. aAb = atypical absence; Ab = absence; At = atonic; FDS = focal dyscognitive seizures; FS = febrile seizures; H = hemiclonic; IS = infantile spasms; MJ = myoclonic jerks; NCS = non-convulsive status epilepticus; SE = status epilepticus; T = tonic; TC = tonic-clonic; DS = Diffuse Slowing; ED = epileptiform discharge; GPFA = Generalised Paroxysmal Fast Activity; GPSW = Generalised Polyspike Wave; GSW = Generalised Spike Wave; Hz = Hertz; MFD = Multi-Focal Discharges; PPR = Photoparoxysmal Response

**Table 3**

Pathogenic variants in known epileptic encephalopathy genes

| Gene | Proband (Gender) | Inheritance (inference) | cDNA change | Protein change | Diagnosis |
|---|---|---|---|---|---|
| *SCN1A* | T23445(F) | *De novo* (P) | c.4836delC | Ile1613Phefs*5 | Dravet |
| | T1639(M) | Segregates (P) | c.5962G>A | Arg1988Trp | Epilepsy-Aphasia, FS + |
| | T18466 (M) | *De novo* (P) | c.4033G>A | Pro1345Ser | EOEE |
| | T18594 (M) | Segregates (P) % | c.133C>T | Asp45Asn | Epilepsy-Aphasia, FS + |
| | T19875 (F) | *De novo* (P) | c.3977G>A | Ala1326Val | Epilepsy-Aphasia |
| | T18775 (M) | Segregates (P) % | c.1076T>G | Asn359Thr | Dravet (ICEGTC) [29] |
| | T18997 (M) | Unk, parents unavailable (LP) | c.1209delA | Phe403Leufs*12 | Dravet |
| | T19963 (M) | *De novo* (P) | c.4453T>C | Asn1485Asp | Dravet |
| *SCN2A* | T20632 (M) | *De novo* (P) | c.408G>T | Met136Ile | EOEE evolving to IS |
| | T21005 (F) | *De novo* (P) | c.2715G>C | Lys905Asn | EE |
| | T22816(F) | Unk, father unavailable (LP) | c.2783T>G | Phe928Cys# | EE |
| | T20340 (F) | *De novo* (P) | c.1154delC | Ile1021Tyrfs*16 | LGS |
| | T24127 (F) | *De novo* (P) | c.5645G>A | Arg1882Gln | EE |
| *PCDH19* | T23579 (F) | X-linked, female restricted (P) | c.1681G>A | Pro561Ser | EE |
| | T23305 (F) | X-linked, female restricted (P) | c.2873C>T | Arg958Gln | LGS |
| *CDKL5* | T20819 (M) | *De novo* (P) | c.464-2A>G | Unk | EOEE |
| | T22724 (M) | Inherited from unaffected mother, X-linked (P) | c.433C>T | His145Tyr | EE |
| | T22954 (F) | *De novo* (P) | c.545T>C | Leu182Pro | EOEE |
| | T897(F) | *De novo* (P) | c.2564C>G | Ser855* | IS |
| | T23057 (M) | *De novo* (P) | c.1926delT | Leu642Argfs*16 | IS |
| | T23951 (M) | *De novo* (P) | c.533G>A | Arg178Gln | EOEE |
| | T23234 (F) | *De novo* (P) | c.620G>A | Gly207Glu | EE |
| | T24139 (M) | Unk, parents unavailable (LP) | c.1926delT | Leu642Argfs*16 | EOEE |
| *STXBP1* | T22595 (M) | *De novo* (P) | c.1154delC | Met387Tyrfs*17 | Ohtahara |
| | T1266 (M) | Unk, mother unavailable (LP) | c.1630G>T | Gly544Cys# | LGS |

| Gene | Proband (Gender) | Inheritance (inference) | cDNA change | Protein change | Diagnosis |
|---|---|---|---|---|---|
| | T23151 (F) | *De novo* (P) | c.125C>T | Ser42Phe | EOEE |
| | T23553 (F) | *De novo* (P) | c.238T>C | Ser80Pro | EE |
| | T23122 (M) | *De novo* (P) | c.568C>T | Arg190Trp | EOEE |
| | T22856 (M) | *De novo* (P) | c.1060T>C | Cys354Arg | Ohtahara |
| | T23289 (M) | *De novo* (P) | c.1708G>A | Thr570Ala | EE |
| *UBE3A* | T23859 (F) | Inherited from unaffected mother, affected sib also mutation positive (P) | c.1585G>A | Arg506Cys& | EE – features suggestive of Angelman syndrome |
| *SCN8A* | T3929 (M) | Inherited from somatic mosaic father @ (P) | c.3868C>G | Leu1290Val | EE |
| *KCNQ2* | T24158 (M) | *De novo* (P) | c.587G>A | Ala196Val | EOEE |
| | T23919 (F) | *De novo* (P) | c.602C>T | Arg201His | EOEE/IS |
| *PNPO* | T23451 (M) | Homozygous recessive (P) | c.686G>A | Arg229Gln^ | EME |
| *PNKP* | T23141 (M) | Homozygous recessive (P) | c.58G>A | Pro20Ser | EE |

M- male F-female; P – pathogenic, LP – likely pathogenic;

%variant segregates with the disorder, pedigrees Supplementary Figure 5;

#Two missense variants likely pathogenic (see methods);

&known pathogenic variant [30],

@father is somatic mosaic, with 13% of cells carrying alternate, pathogenic allele;

^known pathogenic variant dbSNP:rs104894629 [31].

Accession numbers: *SCN1A*, NM_001165963.1, NP_001159435.1; *SCN2A*, NM_021007.2, NP_066287.2; *PCDH19*, NM_001184880.1, NP_001171809.1; *CDKL5*, NM_001037343.1, NP_001032420.1; *STXBP1*, NP_003165.3, NM_ 001032221.3; *UBE3A*, NM_000462.3, NP_000453.2; *SCN8A*, NM_001177984.2; NP_001171455.1; *KCNQ2*, NM_004518.4, NP_004509.2; *PNPO*, NM018129.3, NP_060599.1;*PNKP*, NM_007254.3, NP009185.2

EE = epileptic encephalopathy not otherwise specified: EOEE = early onset epileptic encephalopathy; EME = early myoclonic encephalopathy; FS + = febrile seizures plus; ICEGTC = intractable childhood epilepsy with generalized tonic clonic seizures; LGS = Lennox Gastaut syndrome; IS = Infantile spasms, Unk = unknown